FORM PTO-1390 U.S. DEPARTMENT OF COMMERCE PATENT AND TRADEMARK OFFICE

TRANSMITTAL LETTER TO THE UNITED STATES DESIGNATED/ELECTED OFFICE (DO/EO/US) CONCERNING A FILING UNDER 35 U.S.C. 371

ATTORNEY'S DOCKET NUMBER
PF-0568 USN

U.S. APPLICATION NO. (If known, see 37 CFR 1.5)
TO BE ASSIGNED 09/744314

INTERNATIONAL APPLICATION NO	INTERNATIONAL FILING DATE	PRIORITY DATE CLAIMED
PCT/US99/17167	30 July 1999	31 July 1998

TITLE OF INVENTION
HUMAN CYSTOSKELETAL PROTEINS

APPLICANT(S) FOR DO/EO/US
INCYTE PHARMACEUTICALS, INC.; BANDMAN, Olga; TANG, Y. Tom; YUE, Henry; CORLEY, Neil C.; GUEGLER, Karl J.; AZIMZAI, Yalda; PATTERSON, Chandra; LAL, Preeti; BAUGHN, Mariah R.

Applicant herewith submits to the United States Designated/Elected Office (DO/EO/US) the following items and other information:

1. ☒ This is the **FIRST** submission of items concerning a filing under 35 U.S.C. 371.
2. ☐ This is a **SECOND** or **SUBSEQUENT** submission of items concerning a filing under 35 U.S.C. 371
3. ☐ This is an express request to promptly begin national examination procedures (35 U.S.C. 371 (f)).
4. ☐ The US has been elected by the expiration of 19 months from the priority date (PCT Article 31)
5. ☒ A copy of the International Application as filed (35 U.S.C. 371(c)(2))
   a. ☐ is attached hereto (required only if not communicated by the International Bureau)
   b. ☐ has been communicated by the International Bureau.
   c. ☒ is not required, as the application was filed in the United States Receiving Office (RO/US).
6. ☐ An English language translation of the International Application as filed (35 U.S.C. 371(c)(2)).
7. ☒ Amendments to the claims of the International Application under PCT Article 19 (35 U.S.C. 371(c)(3))
   a. ☐ are attached hereto (required only if not communicated by the International Bureau).
   b. ☐ have been communicated by the International Bureau
   c. ☐ have not been made; however, the time limit for making such amendments has NOT expired.
   d. ☒ have not been made and will not be made.
8. ☐ An English language translation of the amendments to the claims under PCT Article 19 (35 U.S.C. 371(c)(3))
9. ☐ An oath or declaration of the inventor(s) (35 U.S.C. 371(c)(4)).
10. ☐ An English language translation of the annexes to the International Preliminary Examination Report under PCT Article 36 (35 U.S.C. 371(c)(5)).

Items 11 to 16 below concern document(s) or information included:

11. ☐ An Information Disclosure Statement under 37 CFR 1.97 and 1.98.
12. ☐ An assignment document for recording. A separate cover sheet in compliance with 37 CFR 3.27 and 3.31 is included.
13. ☐ A FIRST preliminary amendment.
    ☐ A SECOND or SUBSEQUENT preliminary amendment.
14. ☐ A substitute specification
15. ☐ A change of power of attorney and/or address letter.
16. ☒ Other items or information:

1) Transmittal Letter (2 pp, in duplicate)
2) Return Postcard
3) Express Mail Label No.: **EL 743 380 089 US**

JC07 Rec'd PCT/PTO 18 JAN 2001

U.S. APPLICATION NO. (IF KNOWN, SEE 37 CFR 1.5) TO BE ASSIGNED 09/744314	INTERNATIONAL APPLICATION NO PCT/US99/17167	ATTORNEY'S DOCKET NUMBER PF-0568 USN

17. ☐ The following fees are submitted.

BASIC NATIONAL FEE (37 CFR 1.492(a)(1)-(5)):

Neither international preliminary examination fee (37 CFR 1.482) nor international search fee (37 CFR 1.445(a)(2)) paid to USPTO and International Search Report not prepared by the EPO or JPO.....$1000.00

☐International preliminary examination fee (37 CFR 1.482) not paid to USPTO but International Search Report prepared by the EPO or JPO..$860.00

International preliminary examination fee (37 CFR 1.482) not paid to USPTO but international search fee (37 CFR 1.445(a)(2)) paid to USPTO.. .....$710 00

☒**International preliminary examination fee paid to USPTO (37 CFR 1.482) but all claims did not satisfy provisions of PCT Article 33(1)-(4).....$690.00**

☐International preliminary examination fee paid to USPTO (37 CFR 1.482) and all claims satisfied provisions of PCT Article 33(1)-(4).....$100.00

ENTER APPROPRIATE BASIC FEE AMOUNT =				$690.00	
Surcharge of **$130.00** for furnishing the oath or declaration later than ☐ 20 ☐ 30 months from the earliest claimed priority date (37 CFR 1.492(e))				$	
CLAIMS	NUMBER FILED		NUMBER EXTRA	RATE	
Total Claims	20	=	0	X **$ 18.00**	$
Independent Claims	2	=	0	X **$ 80.00**	$
MULTIPLE DEPENDENT CLAIM(S) (if applicable)				+ **$270.00**	$
TOTAL OF ABOVE CALCULATIONS =				$690.00	
☐ Applicant claims small entity status. See 37 CFR 1.27 The fees indicated above are reduced by 1/2.				$	
SUBTOTAL =				$690.00	
Processing fee of **$130.00** for furnishing the English translation later than ☐ 20 ☐ 30 months from the earliest claimed priority date (37 CFR 1492(f)).				+ $	
TOTAL NATIONAL FEE =				$690.00	
Fee for recording the enclosed assignment (37 CFR 1.21(h)). The assignment must be accompanied by the appropriate cover sheet (37 CFR 3 28, 3.31). **$40.00** per property				+ $	
TOTAL FEES ENCLOSED =				$690 00	
				Amount to be Refunded	$
				Charged	$

a. ☐ A check in the amount of $_____ to cover the above fees is enclosed.
b. ☐ Please charge my Deposit Account No. 09-0108 in the amount of $ 690.00 to cover the above fees.
c. ☒ The Commissioner is hereby authorized to charge any additional fees which may be required, or credit any overpayment to Deposit Account No. 09-0108 A duplicate copy of this sheet is enclosed.

NOTE: Where an appropriate time limit under 37 CFR 1.494 or 1.495 has not been met, a petition to revive (37 CFR 1.137(a) or (b)) must be filed and granted to restore the application to pending status.

SEND ALL CORRESPONDENCE TO:

INCYTE GENOMICS, INC.
3160 Porter Drive
Palo Alto, CA 94304

Diana Hamlet-Cox
SIGNATURE

NAME: Diana Hamlet-Cox

REGISTRATION NUMBER 33,302

DATE: 17 January 2001

HUMAN CYTOSKELETAL PROTEINS

TECHNICAL FIELD

This invention relates to nucleic acid and amino acid sequences of human cytoskeletal proteins and to the use of these sequences in the diagnosis, treatment, and prevention of cell proliferative, immunological, vesicle trafficking, reproductive, smooth muscle, developmental, and nervous disorders.

BACKGROUND OF THE INVENTION

The physical-biochemical processes of cell motility, organelle movement, chromosome movement, cytokinesis, and the generation of cell shape are all dependent on a complex of protein fibers found in the cytoplasm. This protein complex is termed the cytoskeleton. The cytoskeleton of eukaryotic cells has three major filamentous systems. These systems are the actin filaments, intermediate filaments, and microtubules. Each of these filamentous systems is assembled from different proteins, including actin, myosin, tubulins, and intermediate filament proteins. Different cell types and tissues express specific isoforms of the proteins which comprise these filaments. In some cases distinct isoforms and mRNA splice variants are associated with cell-type specific functions. (Lees-Miller, J.P. and Helfman, D.M. (1991) BioEssays 13:429-437.)

The actin filamentous system largely regulates cell motility, in particular generation of muscle tissue contraction and relaxation. The actin filamentous system comprises the thick filament and the thin filament. The thick filament is composed of myosin and the thin filament contains actin and a protein complex of troponin and tropomyosin. Activation of myosin binding to actin is initiated by Ca^{2+}-dependent phosphorylation of myosin light chains by Ca^{2+}-dependent protein kinases. This mechanism is termed the primary Ca^{2+}-dependent mechanism. The interaction between actin and myosin which drives muscle contraction is regulated by binding of Ca^{2+} ions to the troponin-tropomyosin complex and is termed the secondary Ca^{2+}-dependent mechanism. The bound troponin-tropomyosin complex inhibits the interaction between actin and myosin at low cellular concentrations of Ca^{2+} (< 1 μM). Following a nerve generated signal, Ca^{2+} is released from the sarcoplasmic reticulum (SR). At high levels (> 1 μM) Ca^{2+} binds to four sites on troponin and affects the specific molecular interactions between tropomyosin and the actin filament. This reveals the myosin-binding sites on actin, allowing ATP-generated movement of the myosin along the actin filament (contraction). An SR membrane, Ca^{2+}-activated ATPase pumps Ca^{2+} back into the SR thereby returning cytoplasmic Ca^{2+} ion levels to <1 μM. Depletion of cytosolic Ca^{2+} enables the actin-troponin-tropomyosin complex to reform and myosin-actin interactions cease (relaxation). (Reedy, M.K. et al. (1994) Curr. Biol. 4:624-626.)

Different isoforms of tropomyosin have been identified in muscle and non-muscle tissue. The α gene isoform splice variants are in striated and smooth muscle, brain, and fibroblasts; the βchain isoform splice variants are in skeletal muscle and smooth muscle fibroblasts; the hTMnm isoform splice variants are in skeletal muscle and fibroblasts; and the TM-4 isoform splice variant is in rat platelets. (Lees-

Miller, supra; Pittenger, M.F., et al. (1994) Curr. Biol. 6:96-104.)

Secondary Ca^{2+}-dependent mechanisms can modulate the contractile state of the cell. These secondary mechanisms include, but are not limited to, interactions between I) actin, tropomyosin and calponin; ii) actin, myosin, tropomyosin, and caldesmon; iii) actin, tropomyosin, and titin; iv) actin, tropomyosin, and tropomodulin; and v) protein kinase C-dependent chemical modification of actin filament complexes. (Walsh, M.P. (1991) Biochem. Cell. Biol. 69:771-800; Fowler, V.M. (1997) Soc. Gen. Physiol. Ser. 52:79-89.)

Neuronal development and maturation are accompanied by dynamic spatial sorting of tropomyosin isoforms into different cellular compartments. (Gunning, P. et al. (1997) Anat. Embryol. (Berl.) 195:311-315.) Analysis of the developmental changes in the protein compositions of the brain has identified novel developmentally regulated actin-binding proteins termed debrins. Debrin and tropomyosin competitively bind to actin filaments. The exclusion of tropomyosin from actin filaments by debrin results in the appearance of thick, curving bundles of actin filaments and the formation of cell processes in cultured cells. (Shirao, T. (1995) J. Biochem. (Tokyo) 117:231-236.)

Myocardial performance is impaired in chronically diabetic rats. Malhotra and Sanghi have suggested (1997, Cardiovasc. Res. 34:34-40) that diabetes-associated cardiovascular diseases may involve proteins of the actin filamentous system, in particular myosin and troponin. Phosphorylation of troponin has been associated with altered calcium force in isolated muscle preparations. This may be due to changes in the troponin-tropomyosin-actin complex to prevent or reduce interaction(s) with myosin. It was suggested that phosphorylation of troponin could contribute to depressed myocardial contractility in experimental diabetes. (Malhotra and Sanghi, supra.) Mutations in troponin and tropomyosin are associated with familial hypertrophic cardiomyopathy. (Palmiter, K.A. and Solaro, R.J. (1997) Basic. Res. Cardiol. 92 (suppl.1):63-74.)

The term receptor describes proteins that specifically recognize other molecules. The bulk of receptors are cell surface proteins which bind extracellular ligands and lead to cellular responses including growth, differentiation, endocytosis, and immune response. Cell surface receptors are typically integral membrane proteins of the plasma membrane. These receptors recognize compounds, e.g., catecholamine and peptide hormones, growth and differentiation factors, cytokines, small peptide factors, neurotransmitters, and circulatory system-borne signaling molecules. Cell surface receptors on immune system cells recognize antigens, antibodies, and major histocompatibility complex (MHC)-bound peptide. Other cell surface receptors bind ligands to be internalized by the cell. (Lodish, H. et al. (1995) Molecular Cell Biology, Scientific American Books, New York, NY, p. 723; Mikhailenko, I. et al. (1997) J. Biol. Chem. 272:6784-6791.)

The discovery that the transforming oncogene, trk, is a chimeric protein between tropomyosin and the membrane domain and intracellular domain of nerve growth factor receptor has linked mutation of tropomyosin(s) and abnormal expression of tropomyosin(s) to childhood malignancy neuroblastoma.

(Pahlman, S. and Hoehner, J.C. (1996) Mol. Med. Today 2:432-438.) In addition, experimental chimeric proteins containing receptor extracellular and transmembrane domains linked to the integrin β1 cytoplasmic domain can be induced by exogenous factors to cause cytoskeletal reorganization. (Smilenov, L. et al. (1994) Mol. Biol. Cell 5:1215-1223.) These results show that naturally-occurring and synthetic chimeric proteins which combine extracellular, transmembrane, and cytosolic elements of otherwise distinct individual proteins have additional and synergistic roles in tissue biology.

Cell motility is governed by the interaction between cytoskeletal and other cellular proteins. Cytoskeletal proteins which are involved in the generation of motive force within the cell are termed contractile proteins. The energy for this force is generated by ATP.

Two predominant contractile proteins in all animal cells are actin and myosin. Actin is present in both soluble and polymerized forms. For example, filamentous (polymerized) actin interacts with myosin to contract or relax muscle tissues, to transport cell organelles through the intracellular medium, to cause cell movement, and to separate daughter nuclei during cytokinesis.

Multiprotein complexes associate with actin and myosin in vivo. Actin polymerization can be initiated, prevented, or reversed by post-translational protein modification and changes in the constituent proteins of the multiprotein complexes. Examples of multiprotein complex constituent proteins include trychohyalin, p16-Arc, and actinin. Trichohyalin is a cross-linking protein that modulates actin polymerization and functions in intermediate filament- nuclear matrix anchoring (Lee, S.C. (1993) J. Biol. Chem. 268:12164-12176). p16-Arc is a subunit of the human Arp2/3 multiprotein complex. The Arp2/3 complex is localized to the actin-rich lamellipodial protrusions of cells where it is proposed to promote actin assembly and cellular locomotion (Welch, M.D. et al. (1997) J. Cell Biol. 138:375-384). Actinin functions in the linkage of actin to the cell membrane (Honda, K. et al. (1998) J. Cell Biol. 140:1383-1393).

Cytoskeletal proteins are involved in the regulation of muscle contraction. Vertebrate smooth muscle contraction is dependent upon levels of cAMP and intracellular calcium ions (Ca^{2+}). The sarcoplasmic reticulum (SR) serves as an intracellular store of Ca^{2+}. Following hormonal stimulation of the second messenger molecule, inositoltrisphosphate, Ca^{2+} is briefly released from the SR into the surrounding cytoplasm. Ca^{2+} binds to calmodulin (CaM), which activates CaM-dependent myosin light chain protein kinase (MLCK), which then phosphorylates MLC. In relaxed skeletal muscle, myosin is prevented from interacting with actin by binding to tropomyosin. An increase in Ca^{2+} causes a conformational change in tropomyosin-actin binding that leads to the release of actin. This allows actin to interact with phosphorylated MLC forming actinomyosin and initiating the contraction process. Muscle relaxation is brought about by active transport of Ca^{2+} into the SR by a calcium ATPase pump and activation of MLCK by a cAMP-dependent protein kinase (PKA). Interactions between MLCK and PKA may be modulated by other proteins. In particular, telokin, a kinase-related protein encoded by the 3' region of the vertebrate smooth muscle MLCK gene, inhibits MLCK-dependent phosphorylation of MLC

by modulating both the oligomeric state of MLCK and MLCK's interaction with dephosphorylated myosin filaments (Nieznanski, K. and Sobieszek, A. (1997) Biochem. J. 322:65-71). Caldesmon is a protein involved in smooth muscle contraction that performs a role similar to that of tropomyosin in skeletal muscle. Caldesmon forms a complex with tropomyosin and actin that prevents binding of myosin to actin. Phosphorylation of caldesmon by casein kinase II releases its binding to tropomyosin and actin permitting the cross-linking of myosin to actin and the initiation of smooth muscle contraction (Sutherland, C. et al. (1994) J. Muscle Res. Cell. Motil. 15:440-456). Elevation of intracellular cGMP and activation of protein kinase G (PKG) produces relaxation of smooth muscle (Li, H. et al. (1996) J. Vasc. Res. 33:99-110).

Cytoskeletal filament proteins which generate cellular movement are components of flagella and cilia. Flagella and cilia are the hair-like structures which protrude from many cells and are composed of proteinaceous cylinders known as axonemes. The major mass of the axoneme consists of tubulins which polymerize to form microtubules. Nine microtubular doublets typically surround, and are linked to, a central pair of microtubules. Intermediate filament proteins, such as tektins, interact with microtubules to regulate movement. Tektins are predicted to form extended α-helical rods capable of forming coiled-coil structures which are interrupted by short non-helical linkers (Norrander, J.M. et al. (1996) J. Mol. Biol. 257:385-397). Microtubule-associated proteins (MAPs) regulate cell division and cell motility by modulation of microtubule formation.

Cytoskeletal proteins are implicated in several diseases. Pathologies such as muscular dystrophy, nephrotic syndrome, and dilated cardiomyopathy have been associated with differential expression of alpha-actinin-3 (Vainzof, M. et al. (1997) Neuropediatrics 28:223-228; Smoyer, W.E. and Mundel, P. (1998) J. Mol. Med. 76:172-183; and Sussman, M.A. et al. (1998) J. Clin. Invest. 101:51-61). Alpha-actinin and several MAPs are present in Hirano bodies, which are observed more frequently in the elderly and in patients with neurodegenerative diseases such as Alzheimer's disease (Maciver, S.K. and Harrington, C.R. (1995) Neuroreport. 6:1985-1988). Actinin-4, a novel actin-bundling protein, appears to be associated with the cell motility of metastatic cancer cells. Other disease associations include premature chromosome condensation which is frequently observed in dividing cells from tumor tissue (Honda et al. supra; Murnane, J.P. (1995) Cancer Metastasis Rev. 14:17-29) and the significant roles of axonemal and assembly MAPs in viral pathogenesis (Sodeik, B. et al. (1997) J. Cell Biol. 136:1007-1021).

The discovery of new human cytoskeletal proteins and the polynucleotides encoding them satisfies a need in the art by providing new compositions which are useful in the diagnosis, prevention, and treatment of cell proliferative, immunological, vesicle trafficking, reproductive, smooth muscle, developmental, and nervous disorders.

SUMMARY OF THE INVENTION

The invention features substantially purified polypeptides, human cytoskeletal proteins, referred to collectively as "HCYT" and individually as "HCYT-1," "HCYT-2," "HCYT-3," "HCYT-4," "HCYT-5," "HCYT-6," "HCYT-7," and "HCYT-8." In one aspect, the invention provides a substantially purified polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-8 and fragments thereof.

The invention further provides a substantially purified variant having at least 90% amino acid identity to at least one of the amino acid sequences selected from the group consisting of SEQ ID NO:1-8, and fragments thereof. The invention also provides an isolated and purified polynucleotide encoding the polypeptide comprising an amino acid selected from the group consisting of SEQ ID NO:1-8 and fragments thereof. The invention also includes an isolated and purified polynucleotide variant having at least 70% polynucleotide sequence identity to the polynucleotide encoding the polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-8 and fragments thereof.

Additionally, the invention provides an isolated and purified polynucleotide which hybridizes under stringent conditions to the polynucleotide encoding the polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-8 and fragments thereof. The invention also provides an isolated and purified polynucleotide having a sequence which is complementary to the polynucleotide encoding the polypeptide comprising the amino acid sequence selected from the group consisting of SEQ ID NO:1-8 and fragments thereof.

The invention also provides a method for detecting a polynucleotide in a sample containing nucleic acids, the method comprising the steps of (a) hybridizing the complement of the polynucleotide sequence to at least one of the polynucleotides of the sample, thereby forming a hybridization complex; and (b) detecting the hybridization complex, wherein the presence of the hybridization complex correlates with the presence of a polynucleotide in the sample. In one aspect, the method further comprises amplifying the polynucleotide prior to hybridization.

The invention also provides an isolated and purified polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:9-16 and fragments thereof. The invention further provides an isolated and purified polynucleotide variant having at least 70% polynucleotide sequence identity to the polynucleotide sequence selected from the group consisting of SEQ ID NO:9-16, and fragments thereof. The invention also provides an isolated and purified polynucleotide having a sequence which is complementary to the polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:9-16 and fragments thereof.

The invention further provides an expression vector containing at least a fragment of the polynucleotide encoding the polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-8 and fragments thereof. In another aspect, the expression vector is contained within a host cell.

The invention also provides a method for producing a polypeptide, the method comprising the

steps of: (a) culturing the host cell containing an expression vector containing at least a fragment of a polynucleotide under conditions suitable for the expression of the polypeptide; and (b) recovering the polypeptide from the host cell culture.

The invention also provides a pharmaceutical composition comprising a substantially purified polypeptide having the amino acid sequence selected from the group consisting of SEQ ID NO:1-8 and fragments thereof, in conjunction with a suitable pharmaceutical carrier.

The invention further includes a purified antibody which binds to a polypeptide selected from the group consisting of SEQ ID NO:1-8 and fragments thereof. The invention also provides a purified agonist and a purified antagonist to the polypeptide.

The invention also provides a method for treating or preventing a disorder associated with decreased expression or activity of HCYT, the method comprising administering to a subject in need of such treatment an effective amount of a pharmaceutical composition comprising a substantially purified polypeptide having the amino acid sequence selected from the group consisting of SEQ ID NO:1-8 and fragments thereof, in conjunction with a suitable pharmaceutical carrier.

The invention also provides a method for treating or preventing a disorder associated with increased expression or activity of HCYT, the method comprising administering to a subject in need of such treatment an effective amount of an antagonist of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-8 and fragments thereof.

BRIEF DESCRIPTION OF THE TABLES AND FIGURES

Table 1 shows nucleotide and polypeptide sequence identification numbers (SEQ ID NO), clone identification numbers (clone ID), cDNA libraries, and cDNA fragments used to assemble full-length sequences encoding HCYT.

Table 2 shows features of each polypeptide sequence including potential motifs, homologous sequences, and methods and algorithms used for identification of HCYT.

Table 3 shows the tissue-specific expression patterns of each nucleic acid sequence as determined by northern analysis, diseases or disorders associated with these tissues, and the vector into which each cDNA was cloned.

Table 4 describes the tissues used to construct the cDNA libraries from which Incyte cDNA clones encoding HCYT were isolated.

Table 5 shows the programs, their descriptions, references, and threshold parameters used to analyze HCYT.

Figure 1 shows the amino acid sequence alignments between HCYT-8 (2195418; SEQ ID NO:8) and human p16-Arc subunit (GI 2282042; SEQ ID NO:17) produced using the multisequence alignment program of LASERGENE software (DNASTAR, Madison WI).

DESCRIPTION OF THE INVENTION

Before the present proteins, nucleotide sequences, and methods are described, it is understood that this invention is not limited to the particular machines, materials and methods described, as these may vary. It is also to be understood that the terminology used herein is for the purpose of describing particular embodiments only, and is not intended to limit the scope of the present invention which will be limited only by the appended claims.

It must be noted that as used herein and in the appended claims, the singular forms "a," "an," and "the" include plural reference unless the context clearly dictates otherwise. Thus, for example, a reference to "a host cell" includes a plurality of such host cells, and a reference to "an antibody" is a reference to one or more antibodies and equivalents thereof known to those skilled in the art, and so forth.

Unless defined otherwise, all technical and scientific terms used herein have the same meanings as commonly understood by one of ordinary skill in the art to which this invention belongs. Although any machines, materials, and methods similar or equivalent to those described herein can be used to practice or test the present invention, the preferred machines, materials and methods are now described. All publications mentioned herein are cited for the purpose of describing and disclosing the cell lines, protocols, reagents and vectors which are reported in the publications and which might be used in connection with the invention. Nothing herein is to be construed as an admission that the invention is not entitled to antedate such disclosure by virtue of prior invention.

DEFINITIONS

"HCYT" refers to the amino acid sequences of substantially purified HCYT obtained from any species, particularly a mammalian species, including bovine, ovine, porcine, murine, equine, and preferably the human species, from any source, whether natural, synthetic, semi-synthetic, or recombinant.

The term "agonist" refers to a molecule which, when bound to HCYT, increases or prolongs the duration of the effect of HCYT. Agonists may include proteins, nucleic acids, carbohydrates, or any other molecules which bind to and modulate the effect of HCYT.

An "allelic variant" is an alternative form of the gene encoding HCYT. Allelic variants may result from at least one mutation in the nucleic acid sequence and may result in altered mRNAs or in polypeptides whose structure or function may or may not be altered. Any given natural or recombinant gene may have none, one, or many allelic forms. Common mutational changes which give rise to allelic variants are generally ascribed to natural deletions, additions, or substitutions of nucleotides. Each of these types of changes may occur alone, or in combination with the others, one or more times in a given sequence.

"Altered" nucleic acid sequences encoding HCYT include those sequences with deletions, insertions, or substitutions of different nucleotides, resulting in a polynucleotide the same as HCYT or a polypeptide with at least one functional characteristic of HCYT. Included within this definition are polymorphisms which may or may not be readily detectable using a particular oligonucleotide probe of

the polynucleotide encoding HCYT, and improper or unexpected hybridization to allelic variants, with a locus other than the normal chromosomal locus for the polynucleotide sequence encoding HCYT. The encoded protein may also be "altered," and may contain deletions, insertions, or substitutions of amino acid residues which produce a silent change and result in a functionally equivalent HCYT. Deliberate amino acid substitutions may be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature of the residues, as long as the biological or immunological activity of HCYT is retained. For example, negatively charged amino acids may include aspartic acid and glutamic acid, positively charged amino acids may include lysine and arginine, and amino acids with uncharged polar head groups having similar hydrophilicity values may include leucine, isoleucine, and valine; glycine and alanine; asparagine and glutamine; serine and threonine; and phenylalanine and tyrosine.

The terms "amino acid" or "amino acid sequence" refer to an oligopeptide, peptide, polypeptide, or protein sequence, or a fragment of any of these, and to naturally occurring or synthetic molecules. In this context, "fragments," "immunogenic fragments," or "antigenic fragments" refer to fragments of HCYT which are preferably at least 5 to about 15 amino acids in length, most preferably at least 14 amino acids, and which retain some biological activity or immunological activity of HCYT. Where "amino acid sequence" is recited to refer to an amino acid sequence of a naturally occurring protein molecule, "amino acid sequence" and like terms are not meant to limit the amino acid sequence to the complete native amino acid sequence associated with the recited protein molecule.

"Amplification" relates to the production of additional copies of a nucleic acid sequence. Amplification is generally carried out using polymerase chain reaction (PCR) technologies well known in the art.

The term "antagonist" refers to a molecule which, when bound to HCYT, decreases the amount or the duration of the effect of the biological or immunological activity of HCYT. Antagonists may include proteins, nucleic acids, carbohydrates, antibodies, or any other molecules which decrease the effect of HCYT.

The term "antibody" refers to intact molecules as well as to fragments thereof, such as Fab, $F(ab')_2$, and Fv fragments, which are capable of binding the epitopic determinant. Antibodies that bind HCYT polypeptides can be prepared using intact polypeptides or using fragments containing small peptides of interest as the immunizing antigen. The polypeptide or oligopeptide used to immunize an animal (e.g., a mouse, a rat, or a rabbit) can be derived from the translation of RNA, or synthesized chemically, and can be conjugated to a carrier protein if desired. Commonly used carriers that are chemically coupled to peptides include bovine serum albumin, thyroglobulin, and keyhole limpet hemocyanin (KLH). The coupled peptide is then used to immunize the animal.

The term "antigenic determinant" refers to that fragment of a molecule (i.e., an epitope) that makes contact with a particular antibody. When a protein or a fragment of a protein is used to immunize

a host animal, numerous regions of the protein may induce the production of antibodies which bind specifically to antigenic determinants (given regions or three-dimensional structures on the protein). An antigenic determinant may compete with the intact antigen (i.e., the immunogen used to elicit the immune response) for binding to an antibody.

The term "antisense" refers to any composition containing a nucleic acid sequence which is complementary to the "sense" strand of a specific nucleic acid sequence. Antisense molecules may be produced by any method including synthesis or transcription. Once introduced into a cell, the complementary nucleotides combine with natural sequences produced by the cell to form duplexes and to block either transcription or translation. The designation "negative" can refer to the antisense strand, and the designation "positive" can refer to the sense strand.

The term "biologically active," refers to a protein having structural, regulatory, or biochemical functions of a naturally occurring molecule. Likewise, "immunologically active" refers to the capability of the natural, recombinant, or synthetic HCYT, or of any oligopeptide thereof, to induce a specific immune response in appropriate animals or cells and to bind with specific antibodies.

The terms "complementary" or "complementarity" refer to the natural binding of polynucleotides by base pairing. For example, the sequence "5' A-G-T 3'" bonds to the complementary sequence "3' T-C-A 5'." Complementarity between two single-stranded molecules may be "partial," such that only some of the nucleic acids bind, or it may be "complete," such that total complementarity exists between the single stranded molecules. The degree of complementarity between nucleic acid strands has significant effects on the efficiency and strength of the hybridization between the nucleic acid strands. This is of particular importance in amplification reactions, which depend upon binding between nucleic acids strands, and in the design and use of peptide nucleic acid (PNA) molecules.

A "composition comprising a given polynucleotide sequence" or a "composition comprising a given amino acid sequence" refer broadly to any composition containing the given polynucleotide or amino acid sequence. The composition may comprise a dry formulation or an aqueous solution. Compositions comprising polynucleotide sequences encoding HCYT or fragments of HCYT may be employed as hybridization probes. The probes may be stored in freeze-dried form and may be associated with a stabilizing agent such as a carbohydrate. In hybridizations, the probe may be deployed in an aqueous solution containing salts (e.g., NaCl), detergents (e.g., sodium dodecyl sulfate; SDS), and other components (e.g., Denhardt's solution, dry milk, salmon sperm DNA, etc.).

"Consensus sequence"refers to a nucleic acid sequence which has been resequenced to resolve uncalled bases, extended using XL-PCR kit (Perkin-Elmer, Norwalk CT) in the 5' and/or the 3' direction, and resequenced, or which has been assembled from the overlapping sequences of more than one Incyte Clone using a computer program for fragment assembly, such as the GELVIEW Fragment Assembly system (GCG, Madison WI). Some sequences have been both extended and assembled to produce the consensus sequence.

The term "correlates with expression of a polynucleotide" indicates that the detection of the presence of nucleic acids, the same or related to a nucleic acid sequence encoding HCYT, by northern analysis is indicative of the presence of nucleic acids encoding HCYT in a sample, and thereby correlates with expression of the transcript from the polynucleotide encoding HCYT.

A "deletion"refers to a change in the amino acid or nucleotide sequence that results in the absence of one or more amino acid residues or nucleotides.

The term "derivative" refers to the chemical modification of a polypeptide sequence, or a polynucleotide sequence. Chemical modifications of a polynucleotide sequence can include, for example, replacement of hydrogen by an alkyl, acyl, or amino group. A derivative polynucleotide encodes a polypeptide which retains at least one biological or immunological function of the natural molecule. A derivative polypeptide is one modified by glycosylation, pegylation, or any similar process that retains at least one biological or immunological function of the polypeptide from which it was derived.

The term "similarity" refers to a degree of complementarity. There may be partial similarity or complete similarity. The word "identity" may substitute for the word "similarity." A partially complementary sequence that at least partially inhibits an identical sequence from hybridizing to a target nucleic acid is referred to as "substantially similar." The inhibition of hybridization of the completely complementary sequence to the target sequence may be examined using a hybridization assay (Southern or northern blot, solution hybridization, and the like) under conditions of reduced stringency. A substantially similar sequence or hybridization probe will compete for and inhibit the binding of a completely similar (identical) sequence to the target sequence under conditions of reduced stringency. This is not to say that conditions of reduced stringency are such that non-specific binding is permitted, as reduced stringency conditions require that the binding of two sequences to one another be a specific (i.e., a selective) interaction. The absence of non-specific binding may be tested by the use of a second target sequence which lacks even a partial degree of complementarity (e.g., less than about 30% similarity or identity). In the absence of non-specific binding, the substantially similar sequence or probe will not hybridize to the second non-complementary target sequence.

The phrases "percent identity" or "% identity" refer to the percentage of sequence similarity found in a comparison of two or more amino acid or nucleic acid sequences. Percent identity can be determined electronically, e.g., by using the MEGALIGN program (DNASTAR, Madison WI). The MEGALIGN program can create alignments between two or more sequences according to different methods, e.g., the clustal method. (See, e.g., Higgins, D.G. and P.M. Sharp (1988) Gene 73:237-244.) The clustal algorithm groups sequences into clusters by examining the distances between all pairs. The clusters are aligned pairwise and then in groups. The percentage similarity between two amino acid sequences, e.g., sequence A and sequence B, is calculated by dividing the length of sequence A, minus the number of gap residues in sequence A, minus the number of gap residues in sequence B, into the sum of the residue matches between sequence A and sequence B, times one hundred. Gaps of low or of no

similarity between the two amino acid sequences are not included in determining percentage similarity. Percent identity between nucleic acid sequences can also be counted or calculated by other methods known in the art, e.g., the Jotun Hein method. (See, e.g., Hein, J. (1990) Methods Enzymol. 183:626-645.) Identity between sequences can also be determined by other methods known in the art, e.g., by varying hybridization conditions.

"Human artificial chromosomes" (HACs) are linear microchromosomes which may contain DNA sequences of about 6 kb to 10 Mb in size, and which contain all of the elements required for stable mitotic chromosome segregation and maintenance.

The term "humanized antibody" refers to antibody molecules in which the amino acid sequence in the non-antigen binding regions has been altered so that the antibody more closely resembles a human antibody, and still retains its original binding ability.

"Hybridization" refers to any process by which a strand of nucleic acid binds with a complementary strand through base pairing.

The term "hybridization complex" refers to a complex formed between two nucleic acid sequences by virtue of the formation of hydrogen bonds between complementary bases. A hybridization complex may be formed in solution (e.g., C_0t or R_0t analysis) or formed between one nucleic acid sequence present in solution and another nucleic acid sequence immobilized on a solid support (e.g., paper, membranes, filters, chips, pins or glass slides, or any other appropriate substrate to which cells or their nucleic acids have been fixed).

The words "insertion" or "addition" refer to changes in an amino acid or nucleotide sequence resulting in the addition of one or more amino acid residues or nucleotides, respectively, to the sequence found in the naturally occurring molecule.

"Immune response" can refer to conditions associated with inflammation, trauma, immune disorders, or infectious or genetic disease, etc. These conditions can be characterized by expression of various factors, e.g., cytokines, chemokines, and other signaling molecules, which may affect cellular and systemic defense systems.

The term "microarray" refers to an arrangement of distinct polynucleotides on a substrate.

The terms "element" or "array element" in a microarray context, refer to hybridizable polynucleotides arranged on the surface of a substrate.

The term "modulate" refers to a change in the activity of HCYT. For example, modulation may cause an increase or a decrease in protein activity, binding characteristics, or any other biological, functional, or immunological properties of HCYT.

The phrases "nucleic acid" or "nucleic acid sequence," as used herein, refer to a nucleotide, oligonucleotide, polynucleotide, or any fragment thereof. These phrases also refer to DNA or RNA of genomic or synthetic origin which may be single-stranded or double-stranded and may represent the sense or the antisense strand, to peptide nucleic acid (PNA), or to any DNA-like or RNA-like material. In this

context, "fragments" refers to those nucleic acid sequences which, comprise a region of unique polynucleotide sequence that specifically identifies SEQ ID NO:9-16, for example, as distinct from any other sequence in the same genome. For example, a fragment of SEQ ID NO:9-16 is useful in hybridization and amplification technologies and in analogous methods that distinguish SEQ ID NO:9-16 from related polynucleotide sequences. A fragment of SEQ ID NO:9-16 is at least about 15-20 nucleotides in length. The precise length of the fragment of SEQ ID NO:9-16 and the region of SEQ ID NO:9-16 to which the fragment corresponds are routinely determinable by one of ordinary skill in the art based on the intended purpose for the fragment. In some cases, a fragment, when translated, would produce polypeptides retaining some functional characteristic, e.g., antigenicity, or structural domain characteristic, e.g., ATP-binding site, of the full-length polypeptide.

The terms "operably associated" or "operably linked" refer to functionally related nucleic acid sequences. A promoter is operably associated or operably linked with a coding sequence if the promoter controls the translation of the encoded polypeptide. While operably associated or operably linked nucleic acid sequences can be contiguous and in the same reading frame, certain genetic elements, e.g., repressor genes, are not contiguously linked to the sequence encoding the polypeptide but still bind to operator sequences that control expression of the polypeptide.

The term "oligonucleotide" refers to a nucleic acid sequence of at least about 6 nucleotides to 60 nucleotides, preferably about 15 to 30 nucleotides, and most preferably about 20 to 25 nucleotides, which can be used in PCR amplification or in a hybridization assay or microarray. "Oligonucleotide" is substantially equivalent to the terms "amplimer," "primer," "oligomer," and "probe," as these terms are commonly defined in the art.

"Peptide nucleic acid" (PNA) refers to an antisense molecule or anti-gene agent which comprises an oligonucleotide of at least about 5 nucleotides in length linked to a peptide backbone of amino acid residues ending in lysine. The terminal lysine confers solubility to the composition. PNAs preferentially bind complementary single stranded DNA or RNA and stop transcript elongation, and may be pegylated to extend their lifespan in the cell.

The term "sample" is used in its broadest sense. A sample suspected of containing nucleic acids encoding HCYT, or fragments thereof, or HCYT itself, may comprise a bodily fluid; an extract from a cell, chromosome, organelle, or membrane isolated from a cell; a cell; genomic DNA, RNA, or cDNA, in solution or bound to a substrate; a tissue; a tissue print; etc.

The terms "specific binding" or "specifically binding" refer to that interaction between a protein or peptide and an agonist, an antibody, or an antagonist. The interaction is dependent upon the presence of a particular structure of the protein, e.g., the antigenic determinant or epitope, recognized by the binding molecule. For example, if an antibody is specific for epitope "A," the presence of a polypeptide containing the epitope A, or the presence of free unlabeled A, in a reaction containing free labeled A and the antibody will reduce the amount of labeled A that binds to the antibody.

The term "stringent conditions" refers to conditions which permit hybridization between polynucleotides and the claimed polynucleotides. Stringent conditions can be defined by salt concentration, the concentration of organic solvent, e.g., formamide, temperature, and other conditions well known in the art. In particular, stringency can be increased by reducing the concentration of salt, increasing the concentration of formamide, or raising the hybridization temperature.

The term "substantially purified" refers to nucleic acid or amino acid sequences that are removed from their natural environment and are isolated or separated, and are at least about 60% free, preferably about 75% free, and most preferably about 90% free from other components with which they are naturally associated.

A "substitution" refers to the replacement of one or more amino acids or nucleotides by different amino acids or nucleotides, respectively.

"Substrate" refers to any suitable rigid or semi-rigid support including membranes, filters, chips, slides, wafers, fibers, magnetic or nonmagnetic beads, gels, tubing, plates, polymers, microparticles and capillaries. The substrate can have a variety of surface forms, such as wells, trenches, pins, channels and pores, to which polynucleotides or polypeptides are bound.

"Transformation" describes a process by which exogenous DNA enters and changes a recipient cell. Transformation may occur under natural or artificial conditions according to various methods well known in the art, and may rely on any known method for the insertion of foreign nucleic acid sequences into a prokaryotic or eukaryotic host cell. The method for transformation is selected based on the type of host cell being transformed and may include, but is not limited to, viral infection, electroporation, heat shock, lipofection, and particle bombardment. The term "transformed" cells includes stably transformed cells in which the inserted DNA is capable of replication either as an autonomously replicating plasmid or as part of the host chromosome, as well as transiently transformed cells which express the inserted DNA or RNA for limited periods of time.

A "variant" of HCYT polypeptides refers to an amino acid sequence that is altered by one or more amino acid residues. The variant may have "conservative" changes, wherein a substituted amino acid has similar structural or chemical properties (e.g., replacement of leucine with isoleucine). More rarely, a variant may have "nonconservative" changes (e.g., replacement of glycine with tryptophan). Analogous minor variations may also include amino acid deletions or insertions, or both. Guidance in determining which amino acid residues may be substituted, inserted, or deleted without abolishing biological or immunological activity may be found using computer programs well known in the art, for example, LASERGENE software (DNASTAR).

The term "variant," when used in the context of a polynucleotide sequence, may encompass a polynucleotide sequence related to HCYT. This definition may also include, for example, "allelic" (as defined above), "splice," "species," or "polymorphic" variants. A splice variant may have significant identity to a reference molecule, but will generally have a greater or lesser number of polynucleotides due

to alternate splicing of exons during mRNA processing. The corresponding polypeptide may possess additional functional domains or an absence of domains. Species variants are polynucleotide sequences that vary from one species to another. The resulting polypeptides generally will have significant amino acid identity relative to each other. A polymorphic variant is a variation in the polynucleotide sequence of a particular gene between individuals of a given species. Polymorphic variants also may encompass "single nucleotide polymorphisms" (SNPs) in which the polynucleotide sequence varies by one base. The presence of SNPs may be indicative of, for example, a certain population, a disease state, or a propensity for a disease state.

THE INVENTION

The invention is based on the discovery of new human cytoskeletal proteins (HCYT), the polynucleotides encoding HCYT, and the use of these compositions for the diagnosis, treatment, or prevention of cell proliferative, immunological, vesicle trafficking, reproductive, smooth muscle, developmental, and nervous disorders.

Table 1 lists the Incyte Clones used to derive full length nucleotide sequences encoding HCYT. Columns 1 and 2 show the sequence identification numbers (SEQ ID NO) of the amino acid and nucleic acid sequences, respectively. Column 3 shows the Clone ID of the Incyte Clone in which nucleic acids encoding each HCYT was identified, and column 4, the cDNA libraries from which these clones were isolated. Column 5 shows Incyte clones, their corresponding cDNA libraries, and shotgun sequences useful as fragments in hybridization technologies, and which are part of the consensus nucleotide sequence of each HCYT.

The columns of Table 2 show various properties of the polypeptides of the invention: column 1 references the SEQ ID NO; column 2 shows the number of amino acid residues in each polypeptide; column 3, potential phosphorylation sites; column 4, potential glycosylation sites; column 5, the amino acid residues comprising signature sequences and motifs; column 6, the identity of each protein; and column 7, analytical methods used to identify each protein through sequence homology and protein motifs. SEQ ID NO:10, which encodes HCYT-2, is a splice variant of SEQ ID NO:9, which encodes HCYT-1. In particular, the nucleotide sequence of SEQ ID NO:10 from nt 190 to nt 811 is identical to the nucleotide sequence of SEQ ID NO:9 from nt 319 to nt 940. The N-terminus of HCYT-1 has two leucine zipper patterns and five additional potential phosphorylation sites. HCYT-1, HCYT-2, HCYT-3, HCYT-4, HCYT-5, and HCYT-6 have chemical and structural similarity with tropomyosin isoforms. The N-termini of HCYT-5 (SEQ ID NO:5) and HCYT-6 (SEQ ID NO:6) have homology to catecholamine receptors and tektins, respectively; the N-terminus of HCYT-4 (SEQ ID NO:4) and the intervening regions of HCYT-5 and HCYT-6 have homology to tropomyosin isoforms; the C-terminus of HCYT-4 has homology to proteins which bind nucleotide di- or triphosphate molecules; and the C-termini of HCYT-5 and HCYT-6 have homology to receptors. SEQ ID NO:7 has various properties that are related to intermediate filament proteins including numerous potential phosphorylation sites, several

leucine zipper motifs, and a tektin signature sequence.

The columns of Table 3 show the tissue-specificity and disease-association of nucleotide sequences encoding HCYT. The first column of Table 3 lists the polynucleotide sequence identifiers. The second column lists tissue categories which express HCYT as a fraction of total tissues expressing HCYT. The third column lists the diseases, disorders, or conditions associated with those tissues expressing HCYT as a fraction of total tissues expressing HCYT. The fourth column lists the vectors used to subclone the cDNA library.

Figure 1 shows that chemical and structural homology, in the context of sequences and motifs, exists between HCYT-8 (SEQ ID NO:8) and human p16-Arc (GI 2282042: SEQ ID NO:17). In particular, the two proteins share 66% identity, the potential phosphorylation sites at S7 and T148, and the potential glycosylation site at N122 in HCYT-8.

The following represent selected fragments of the nucleotide sequences encoding HCYT which are useful as hybridization probes: the fragment of SEQ ID NO:15 from about nucleotide 7052 to about nucleotide 7111; and the fragment of SEQ ID NO:16 from about nucleotide 182 to about nucleotide 235.

The invention also encompasses HCYT variants. A preferred HCYT variant is one which has at least about 80%, more preferably at least about 90%, and most preferably at least about 95% amino acid sequence identity to the HCYT amino acid sequence, and which contains at least one functional or structural characteristic of HCYT.

The invention also encompasses polynucleotides which encode HCYT. In a particular embodiment, the invention encompasses a polynucleotide sequence comprising a sequence selected from the group consisting of SEQ ID NO:9-16, which encodes HCYT.

The invention also encompasses a variant of a polynucleotide sequence encoding HCYT. In particular, such a variant polynucleotide sequence will have at least about 70%, more preferably at least about 85%, and most preferably at least about 95% polynucleotide sequence identity to the polynucleotide sequence encoding HCYT. A particular aspect of the invention encompasses a variant of a polynucleotide sequence comprising a sequence selected from the group consisting of SEQ ID NO:9-16 which has at least about 70%, more preferably at least about 85%, and most preferably at least about 95% polynucleotide sequence identity to a nucleic acid sequence selected from the group consisting of SEQ ID NO:9-16. Any one of the polynucleotide variants described above can encode an amino acid sequence which contains at least one functional or structural characteristic of HCYT.

It will be appreciated by those skilled in the art that as a result of the degeneracy of the genetic code, a multitude of polynucleotide sequences encoding HCYT, some bearing minimal similarity to the polynucleotide sequences of any known and naturally occurring gene, may be produced. Thus, the invention contemplates each and every possible variation of polynucleotide sequence that could be made by selecting combinations based on possible codon choices. These combinations are made in accordance

with the standard triplet genetic code as applied to the polynucleotide sequence of naturally occurring HCYT, and all such variations are to be considered as being specifically disclosed.

Although nucleotide sequences which encode HCYT and its variants are preferably capable of hybridizing to the nucleotide sequence of the naturally occurring HCYT under appropriately selected conditions of stringency, it may be advantageous to produce nucleotide sequences encoding HCYT or its derivatives possessing a substantially different codon usage, e.g., inclusion of non-naturally occurring codons. Codons may be selected to increase the rate at which expression of the peptide occurs in a particular prokaryotic or eukaryotic host in accordance with the frequency with which particular codons are utilized by the host. Other reasons for substantially altering the nucleotide sequence encoding HCYT and its derivatives without altering the encoded amino acid sequences include the production of RNA transcripts having more desirable properties, such as a greater half-life, than transcripts produced from the naturally occurring sequence.

The invention also encompasses production of DNA sequences which encode HCYT and HCYT derivatives, or fragments thereof, entirely by synthetic chemistry. After production, the synthetic sequence may be inserted into any of the many available expression vectors and cell systems using reagents well known in the art. Moreover, synthetic chemistry may be used to introduce mutations into a sequence encoding HCYT or any fragment thereof.

Also encompassed by the invention are polynucleotide sequences that are capable of hybridizing to the claimed polynucleotide sequences, and, in particular, to those shown in SEQ ID NO:9-16 and fragments thereof under various conditions of stringency. (See, e.g., Wahl, G.M. and S.L. Berger (1987) Methods Enzymol. 152:399-407; Kimmel, A.R. (1987) Methods Enzymol. 152:507-511.) For example, stringent salt concentration will ordinarily be less than about 750 mM NaCl and 75 mM trisodium citrate, preferably less than about 500 mM NaCl and 50 mM trisodium citrate, and most preferably less than about 250 mM NaCl and 25 mM trisodium citrate. Low stringency hybridization can be obtained in the absence of organic solvent, e.g., formamide, while high stringency hybridization can be obtained in the presence of at least about 35% formamide, and most preferably at least about 50% formamide. Stringent temperature conditions will ordinarily include temperatures of at least about 30°C, more preferably of at least about 37°C, and most preferably of at least about 42°C. Varying additional parameters, such as hybridization time, the concentration of detergent, e.g., sodium dodecyl sulfate (SDS), and the inclusion or exclusion of carrier DNA, are well known to those skilled in the art. Various levels of stringency are accomplished by combining these various conditions as needed. In a preferred embodiment, hybridization will occur at 30°C in 750 mM NaCl, 75 mM trisodium citrate, and 1% SDS. In a more preferred embodiment, hybridization will occur at 37°C in 500 mM NaCl, 50 mM trisodium citrate, 1% SDS, 35% formamide, and 100 μg/ml denatured salmon sperm DNA (ssDNA). In a most preferred embodiment, hybridization will occur at 42°C in 250 mM NaCl, 25 mM trisodium citrate, 1% SDS, 50 % formamide, and 200 μg/ml ssDNA. Useful variations on these conditions will be readily apparent to those skilled in

the art.

The washing steps which follow hybridization can also vary in stringency. Wash stringency conditions can be defined by salt concentration and by temperature. As above, wash stringency can be increased by decreasing salt concentration or by increasing temperature. For example, stringent salt concentration for the wash steps will preferably be less than about 30 mM NaCl and 3 mM trisodium citrate, and most preferably less than about 15 mM NaCl and 1.5 mM trisodium citrate. Stringent temperature conditions for the wash steps will ordinarily include temperature of at least about 25°C, more preferably of at least about 42°C, and most preferably of at least about 68°C. In a preferred embodiment, wash steps will occur at 25°C in 30 mM NaCl, 3 mM trisodium citrate, and 0.1% SDS. In a more preferred embodiment, wash steps will occur at 42°C in 15 mM NaCl, 1.5 mM trisodium citrate, and 0.1% SDS. In a most preferred embodiment, wash steps will occur at 68°C in 15 mM NaCl, 1.5 mM trisodium citrate, and 0.1% SDS. Additional variations on these conditions will be readily apparent to those skilled in the art.

Methods for DNA sequencing are well known in the art and may be used to practice any of the embodiments of the invention. The methods may employ such enzymes as the Klenow fragment of DNA polymerase I, SEQUENASE (US Biochemical, Cleveland OH), Taq polymerase (Perkin-Elmer), thermostable T7 polymerase (Amersham Pharmacia Biotech, Piscataway NJ), or combinations of polymerases and proofreading exonucleases such as those found in the ELONGASE amplification system (Life Technologies, Gaithersburg MD). Preferably, sequence preparation is automated with machines such as the Robbins HYDRA microdispenser (Robbins Scientific, Sunnyvale CA), Hamilton MICROLAB 2200 (Hamilton, Reno NV), Peltier Thermal Cycler 200 (PTC200; MJ Research, Watertown MA) and the ABI CATALYST 800 (Perkin-Elmer). Sequencing is then carried out using either ABI 373 or 377 DNA Sequencing Systems (Perkin-Elmer) or the MEGABACE 1000 DNA sequencing system (Molecular Dynamics, Sunnyvale CA). The resulting sequences are analyzed using a variety of algorithms which are well known in the art. (See, e.g., Ausubel, F.M. (1997) Short Protocols in Molecular Biology, John Wiley & Sons, New York NY, unit 7.7; Meyers, R.A. (1995) Molecular Biology and Biotechnology, Wiley VCH, New York NY, pp. 856-853.)

The nucleic acid sequences encoding HCYT may be extended utilizing a partial nucleotide sequence and employing various PCR-based methods known in the art to detect upstream sequences, such as promoters and regulatory elements. For example, one method which may be employed, restriction-site PCR, uses universal and nested primers to amplify unknown sequence from genomic DNA within a cloning vector. (See, e.g., Sarkar, G. (1993) PCR Methods Applic. 2:318-322.) Another method, inverse PCR, uses primers that extend in divergent directions to amplify unknown sequence from a circularized template. The template is derived from restriction fragments comprising a known genomic locus and surrounding sequences. (See, e.g., Triglia, T. et al. (1988) Nucleic Acids Res. 16:8186.) A third method, capture PCR, involves PCR amplification of DNA fragments adjacent to known sequences in human and

yeast artificial chromosome DNA. (See, e.g., Lagerstrom, M. et al. (1991) PCR Methods Applic. 1:111-119.) In this method, multiple restriction enzyme digestions and ligations may be used to insert an engineered double-stranded sequence into a region of unknown sequence before performing PCR. Other methods which may be used to retrieve unknown sequences are known in the art. (See, e.g., Parker, J.D. et al. (1991) Nucleic Acids Res. 19:3055-306). Additionally, one may use PCR, nested primers, and PROMOTERFINDER libraries (Clontech, Palo Alto CA) to walk genomic DNA. This procedure avoids the need to screen libraries and is useful in finding intron/exon junctions. For all PCR-based methods, primers may be designed using commercially available software, such as OLIGO 4.06 Primer Analysis software (National Biosciences, Plymouth MN) or another appropriate program, to be about 22 to 30 nucleotides in length, to have a GC content of about 50% or more, and to anneal to the template at temperatures of about 68°C to 72°C.

When screening for full-length cDNAs, it is preferable to use libraries that have been size-selected to include larger cDNAs. In addition, random-primed libraries, which often include sequences containing the 5' regions of genes, are preferable for situations in which an oligo d(T) library does not yield a full-length cDNA. Genomic libraries may be useful for extension of sequence into 5' non-transcribed regulatory regions.

Capillary electrophoresis systems which are commercially available may be used to analyze the size or confirm the nucleotide sequence of sequencing or PCR products. In particular, capillary sequencing may employ flowable polymers for electrophoretic separation, four different nucleotide-specific, laser-stimulated fluorescent dyes, and a charge coupled device camera for detection of the emitted wavelengths. Output/light intensity may be converted to electrical signal using appropriate software (e.g., GENOTYPER and SEQUENCE NAVIGATOR, Perkin-Elmer), and the entire process from loading of samples to computer analysis and electronic data display may be computer controlled. Capillary electrophoresis is especially preferable for sequencing small DNA fragments which may be present in limited amounts in a particular sample.

In another embodiment of the invention, polynucleotide sequences or fragments thereof which encode HCYT may be cloned in recombinant DNA molecules that direct expression of HCYT, or fragments or functional equivalents thereof, in appropriate host cells. Due to the inherent degeneracy of the genetic code, other DNA sequences which encode substantially the same or a functionally equivalent amino acid sequence may be produced and used to express HCYT.

The nucleotide sequences of the present invention can be engineered using methods generally known in the art in order to alter HCYT-encoding sequences for a variety of purposes including, but not limited to, modification of the cloning, processing, and/or expression of the gene product. DNA shuffling by random fragmentation and PCR reassembly of gene fragments and synthetic oligonucleotides may be used to engineer the nucleotide sequences. For example, oligonucleotide-mediated site-directed mutagenesis may be used to introduce mutations that create new restriction sites, alter glycosylation

patterns, change codon preference, produce splice variants, and so forth.

In another embodiment, sequences encoding HCYT may be synthesized, in whole or in part, using chemical methods well known in the art. (See, e.g., Caruthers, M.H. et al. (1980) Nucl. Acids Res. Symp. Ser. 215-223, and Horn, T. et al. (1980) Nucl. Acids Res. Symp. Ser. 225-232.) Alternatively, HCYT itself or a fragment thereof may be synthesized using chemical methods. For example, peptide synthesis can be performed using various solid-phase techniques. (See, e.g., Roberge, J.Y. et al. (1995) Science 269:202-204.) Automated synthesis may be achieved using the ABI 431A Peptide Synthesizer (Perkin-Elmer). Additionally, the amino acid sequence of HCYT, or any part thereof, may be altered during direct synthesis and/or combined with sequences from other proteins, or any part thereof, to produce a variant polypeptide.

The peptide may be substantially purified by preparative high performance liquid chromatography. (See, e.g, Chiez, R.M. and F.Z. Regnier (1990) Methods Enzymol. 182:392-421.) The composition of the synthetic peptides may be confirmed by amino acid analysis or by sequencing. (See, e.g., Creighton, T. (1984) Proteins, Structures and Molecular Properties, WH Freeman, New York NY.)

In order to express a biologically active HCYT, the nucleotide sequences encoding HCYT or derivatives thereof may be inserted into an appropriate expression vector, i.e., a vector which contains the necessary elements for transcriptional and translational control of the inserted coding sequence in a suitable host. These elements include regulatory sequences, such as enhancers, constitutive and inducible promoters, and 5' and 3' untranslated regions in the vector and in polynucleotide sequences encoding HCYT. Such elements may vary in their strength and specificity. Specific initiation signals may also be used to achieve more efficient translation of sequences encoding HCYT. Such signals include the ATG initiation codon and adjacent sequences, e.g. the Kozak sequence. In cases where sequences encoding HCYT and its initiation codon and upstream regulatory sequences are inserted into the appropriate expression vector, no additional transcriptional or translational control signals may be needed. However, in cases where only coding sequence, or a fragment thereof, is inserted, exogenous translational control signals including an in-frame ATG initiation codon should be provided by the vector. Exogenous translational elements and initiation codons may be of various origins, both natural and synthetic. The efficiency of expression may be enhanced by the inclusion of enhancers appropriate for the particular host cell system used. (See, e.g., Scharf, D. et al. (1994) Results Probl. Cell Differ. 20:125-162.)

Methods which are well known to those skilled in the art may be used to construct expression vectors containing sequences encoding HCYT and appropriate transcriptional and translational control elements. These methods include in vitro recombinant DNA techniques, synthetic techniques, and in vivo genetic recombination. (See, e.g., Sambrook, J. et al. (1989) Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Press, Plainview NY, ch. 4, 8, and 16-17; Ausubel, F.M. et al. (1995) Current Protocols in Molecular Biology, John Wiley & Sons, New York NY, ch. 9, 13, and 16.)

A variety of expression vector/host systems may be utilized to contain and express sequences

encoding HCYT. These include, but are not limited to, microorganisms such as bacteria transformed with recombinant bacteriophage, plasmid, or cosmid DNA expression vectors; yeast transformed with yeast expression vectors; insect cell systems infected with viral expression vectors (e.g., baculovirus); plant cell systems transformed with viral expression vectors (e.g., cauliflower mosaic virus, CaMV, or tobacco mosaic virus, TMV) or with bacterial expression vectors (e.g., Ti or pBR322 plasmids); or animal cell systems. The invention is not limited by the host cell employed.

In bacterial systems, a number of cloning and expression vectors may be selected depending upon the use intended for polynucleotide sequences encoding HCYT. For example, routine cloning, subcloning, and propagation of polynucleotide sequences encoding HCYT can be achieved using a multifunctional E. coli vector such as PBLUESCRIPT (Stratagene, La Jolla CA) or pSPORT1 plasmid (Life Technologies). Ligation of sequences encoding HCYT into the vector's multiple cloning site disrupts the *lacZ* gene, allowing a colorimetric screening procedure for identification of transformed bacteria containing recombinant molecules. In addition, these vectors may be useful for in vitro transcription, dideoxy sequencing, single strand rescue with helper phage, and creation of nested deletions in the cloned sequence. (See, e.g., Van Heeke, G. and S.M. Schuster (1989) J. Biol. Chem. 264:5503-5509.) When large quantities of HCYT are needed, e.g. for the production of antibodies, vectors which direct high level expression of HCYT may be used. For example, vectors containing the strong, inducible T5 or T7 bacteriophage promoter may be used.

Yeast expression systems may be used for production of HCYT. A number of vectors containing constitutive or inducible promoters, such as alpha factor, alcohol oxidase, and PGH, may be used in the yeast Saccharomyces cerevisiae or Pichia pastoris. In addition, such vectors direct either the secretion or intracellular retention of expressed proteins and enable integration of foreign sequences into the host genome for stable propagation. (See, e.g., Ausubel, 1995, supra; Grant et al. (1987) Methods Enzymol. 153:516-54; and Scorer, C. A. et al. (1994) Bio/Technology 12:181-184.)

Plant systems may also be used for expression of HCYT. Transcription of sequences encoding HCYT may be driven viral promoters, e.g., the 35S and 19S promoters of CaMV used alone or in combination with the omega leader sequence from TMV (Takamatsu, N. (1987) EMBO J. 6:307-311). Alternatively, plant promoters such as the small subunit of RUBISCO or heat shock promoters may be used. (See, e.g., Coruzzi, G. et al. (1984) EMBO J. 3:1671-1680; Broglie, R. et al. (1984) Science 224:838-843; and Winter, J. et al. (1991) Results Probl. Cell Differ. 17:85-105.) These constructs can be introduced into plant cells by direct DNA transformation or pathogen-mediated transfection. (See, e.g., The McGraw Hill Yearbook of Science and Technology (1992) McGraw Hill, New York NY, pp. 191-196.)

In mammalian cells, a number of viral-based expression systems may be utilized. In cases where an adenovirus is used as an expression vector, sequences encoding HCYT may be ligated into an adenovirus transcription/translation complex consisting of the late promoter and tripartite leader

sequence. Insertion in a non-essential E1 or E3 region of the viral genome may be used to obtain infective virus which expresses HCYT in host cells. (See, e.g., Logan, J. and T. Shenk (1984) Proc. Natl. Acad. Sci. 81:3655-3659.) In addition, transcription enhancers, such as the Rous sarcoma virus (RSV) enhancer, may be used to increase expression in mammalian host cells. SV40 or EBV-based vectors may also be used for high-level protein expression.

Human artificial chromosomes (HACs) may also be employed to deliver larger fragments of DNA than can be contained in and expressed from a plasmid. HACs of about 6 kb to 10 Mb are constructed and delivered via conventional delivery methods (liposomes, polycationic amino polymers, or vesicles) for therapeutic purposes. (See, e.g., Harrington, J.J. et al. (1997) Nat Genet. 15:345-355.)

For long term production of recombinant proteins in mammalian systems, stable expression of HCYT in cell lines is preferred. For example, sequences encoding HCYT can be transformed into cell lines using expression vectors which may contain viral origins of replication and/or endogenous expression elements and a selectable marker gene on the same or on a separate vector. Following the introduction of the vector, cells may be allowed to grow for about 1 to 2 days in enriched media before being switched to selective media. The purpose of the selectable marker is to confer resistance to a selective agent, and its presence allows growth and recovery of cells which successfully express the introduced sequences. Resistant clones of stably transformed cells may be propagated using tissue culture techniques appropriate to the cell type.

Any number of selection systems may be used to recover transformed cell lines. These include, but are not limited to, the herpes simplex virus thymidine kinase and adenine phosphoribosyltransferase genes, for use in *tk* or *apr* cells, respectively. (See, e.g., Wigler, M. et al. (1977) Cell 11:223-232; Lowy, I. et al. (1980) Cell 22:817-823.) Also, antimetabolite, antibiotic, or herbicide resistance can be used as the basis for selection. For example, *dhfr* confers resistance to methotrexate; *neo* confers resistance to the aminoglycosides neomycin and G-418; and *als* or *pat* confer resistance to chlorsulfuron and phosphinotricin acetyltransferase, respectively. (See, e.g., Wigler, M. et al. (1980) Proc. Natl. Acad. Sci. 77:3567-3570; Colbere-Garapin, F. et al. (1981) J. Mol. Biol. 150:1-14.) Additional selectable genes have been described, e.g., *trpB* and *hisD*, which alter cellular requirements for metabolites. (See, e.g., Hartman, S.C. and R.C. Mulligan (1988) Proc. Natl. Acad. Sci. 85:8047-8051.) Visible markers, e.g., anthocyanins, green fluorescent proteins (GFP; Clontech), ß glucuronidase and its substrate ß-glucuronide, or luciferase and its substrate luciferin may be used. These markers can be used not only to identify transformants, but also to quantify the amount of transient or stable protein expression attributable to a specific vector system. (See, e.g., Rhodes, C.A. (1995) Methods Mol. Biol. 55:121-131.)

Although the presence/absence of marker gene expression suggests that the gene of interest is also present, the presence and expression of the gene may need to be confirmed. For example, if the sequence encoding HCYT is inserted within a marker gene sequence, transformed cells containing sequences encoding HCYT can be identified by the absence of marker gene function. Alternatively, a

marker gene can be placed in tandem with a sequence encoding HCYT under the control of a single promoter. Expression of the marker gene in response to induction or selection usually indicates expression of the tandem gene as well.

In general, host cells that contain the nucleic acid sequence encoding HCYT and that express HCYT may be identified by a variety of procedures known to those of skill in the art. These procedures include, but are not limited to, DNA-DNA or DNA-RNA hybridizations, PCR amplification, and protein bioassay or immunoassay techniques which include membrane, solution, or chip based technologies for the detection and/or quantification of nucleic acid or protein sequences.

Immunological methods for detecting and measuring the expression of HCYT using either specific polyclonal or monoclonal antibodies are known in the art. Examples of such techniques include enzyme-linked immunosorbent assays (ELISAs), radioimmunoassays (RIAs), and fluorescence activated cell sorting (FACS). A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering epitopes on HCYT is preferred, but a competitive binding assay may be employed. These and other assays are well known in the art. (See, e.g., Hampton, R. et al. (1990) Serological Methods, a Laboratory Manual, APS Press, St Paul MN, Sect. IV; Coligan, J. E. et al. (1997) Current Protocols in Immunology, Greene Pub. Associates and Wiley-Interscience, New York NY; and Pound, J.D. (1998) Immunochemical Protocols, Humana Press, Totowa NJ).

A wide variety of labels and conjugation techniques are known by those skilled in the art and may be used in various nucleic acid and amino acid assays. Means for producing labeled hybridization or PCR probes for detecting sequences related to polynucleotides encoding HCYT include oligolabeling, nick translation, end-labeling, or PCR amplification using a labeled nucleotide. Alternatively, the sequences encoding HCYT, or any fragments thereof, may be cloned into a vector for the production of an mRNA probe. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes in vitro by addition of an appropriate RNA polymerase such as T7, T3, or SP6 and labeled nucleotides. These procedures may be conducted using a variety of commercially available kits, such as those provided by Amersham Pharmacia Biotech, Promega (Madison WI), and US Biochemical. Suitable reporter molecules or labels which may be used for ease of detection include radionuclides, enzymes, fluorescent, chemiluminescent, or chromogenic agents, as well as substrates, cofactors, inhibitors, magnetic particles, and the like.

Host cells transformed with nucleotide sequences encoding HCYT may be cultured under conditions suitable for the expression and recovery of the protein from cell culture. The protein produced by a transformed cell may be secreted or retained intracellularly depending on the sequence and/or the vector used. As will be understood by those of skill in the art, expression vectors containing polynucleotides which encode HCYT may be designed to contain signal sequences which direct secretion of HCYT through a prokaryotic or eukaryotic cell membrane.

In addition, a host cell strain may be chosen for its ability to modulate expression of the inserted

sequences or to process the expressed protein in the desired fashion. Such modifications of the polypeptide include, but are not limited to, acetylation, carboxylation, glycosylation, phosphorylation, lipidation, and acylation. Post-translational processing which cleaves a "prepro" form of the protein may also be used to specify protein targeting, folding, and/or activity. Different host cells which have specific cellular machinery and characteristic mechanisms for post-translational activities (e.g., CHO, HeLa, MDCK, HEK293, and WI38), are available from the American Type Culture Collection (ATCC, Bethesda MD) and may be chosen to ensure the correct modification and processing of the foreign protein.

In another embodiment of the invention, natural, modified, or recombinant nucleic acid sequences encoding HCYT may be ligated to a heterologous sequence resulting in translation of a fusion protein in any of the aforementioned host systems. For example, a chimeric HCYT protein containing a heterologous moiety that can be recognized by a commercially available antibody may facilitate the screening of peptide libraries for inhibitors of HCYT activity. Heterologous protein and peptide moieties may also facilitate purification of fusion proteins using commercially available affinity matrices. Such moieties include, but are not limited to, glutathione S-transferase (GST), maltose binding protein (MBP), thioredoxin (Trx), calmodulin binding peptide (CBP), 6-His, FLAG, *c-myc*, and hemagglutinin (HA). GST, MBP, Trx, CBP, and 6-His enable purification of their cognate fusion proteins on immobilized glutathione, maltose, phenylarsine oxide, calmodulin, and metal-chelate resins, respectively. FLAG, *c-myc*, and hemagglutinin (HA) enable immunoaffinity purification of fusion proteins using commercially available monoclonal and polyclonal antibodies that specifically recognize these epitope tags. A fusion protein may also be engineered to contain a proteolytic cleavage site located between the HCYT encoding sequence and the heterologous protein sequence, so that HCYT may be cleaved away from the heterologous moiety following purification. Methods for fusion protein expression and purification are discussed in Ausubel (1995, supra, ch 10). A variety of commercially available kits may also be used to facilitate expression and purification of fusion proteins.

In a further embodiment of the invention, synthesis of radiolabeled HCYT may be achieved in vitro using the TNT rabbit reticulocyte lysate or wheat germ extract systems (Promega). These systems couple transcription and translation of protein-coding sequences operably associated with the T7, T3, or SP6 promoters. Translation takes place in the presence of a radiolabeled amino acid precursor, preferably ^{35}S-methionine.

Fragments of HCYT may be produced not only by recombinant production, but also by direct peptide synthesis using solid-phase techniques. (See, e.g., Creighton, supra pp. 55-60.) Protein synthesis may be performed by manual techniques or by automation. Automated synthesis may be achieved, for example, using the ABI 431A Peptide Synthesizer (Perkin-Elmer). Various fragments of HCYT may be synthesized separately and then combined to produce the full length molecule.

THERAPEUTICS

Chemical and structural similarity, e.g., in the context of sequences and motifs, exists between HCYT and human cytoskeletal proteins. In addition, HCYT is expressed in tissues associated with cancer, cell proliferation, fetal development, and inflammation and the immune response, as well as in reproductive, nervous, cardiovascular, developmental, and gastrointestinal tissues. Therefore, HCYT appears to be involved with cell proliferative, immunological, vesicle trafficking, reproductive, smooth muscle, developmental, and nervous disorders. In disorders associated with decreased expression or activity of HCYT, it is desirable to increase the expression or activity of HCYT. In disorders associated with increased expression or activity of HCYT, it is desirable to decrease the expression or activity of HCYT.

Therefore, in one embodiment, HCYT or a fragment or derivative thereof may be administered to a subject to treat or prevent a disorder in which the expression or activity of HCYT is decreased. Examples of such disorders include, but are not limited to, a cell proliferative disorder such as actinic keratosis, arteriosclerosis, atherosclerosis, bursitis, cirrhosis, hepatitis, mixed connective tissue disease (MCTD), myelofibrosis, paroxysmal nocturnal hemoglobinuria, polycythemia vera, psoriasis, primary thrombocythemia; cancers including adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in particular, cancers of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus; an immunological disorder such as actinic keratosis, acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, arteriosclerosis, asthma, atherosclerosis, autoimmune hemolytic anemia, autoimmune thyroiditis, bronchitis, bursitis, cholecystitis, cirrhosis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, paroxysmal nocturnal hemoglobinuria, hepatitis, hypereosinophilia, irritable bowel syndrome, episodic lymphopenia with lymphocytotoxins, mixed connective tissue disease (MCTD), multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, myelofibrosis, osteoarthritis, osteoporosis, pancreatitis, polycythemia vera, polymyositis, psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome, systemic anaphylaxis, systemic lupus erythematosus, systemic sclerosis, primary thrombocythemia, thrombocytopenic purpura, ulcerative colitis, uveitis, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal circulation, viral, bacterial, fungal, parasitic, protozoal, and helminthic infections, and trauma; a vesicle trafficking disorder such as cystic fibrosis, glucose-galactose malabsorption syndrome, hypercholesterolemia, diabetes mellitus, diabetes insipidus, hyper- and hypoglycemia, Grave's disease, goiter, Cushing's disease, and Addison's disease; gastrointestinal disorders including ulcerative colitis, gastric and duodenal ulcers; other conditions associated with abnormal vesicle trafficking, including acquired immunodeficiency syndrome (AIDS);

allergies including hay fever, asthma, and urticaria (hives); autoimmune hemolytic anemia; proliferative glomerulonephritis; inflammatory bowel disease; multiple sclerosis; myasthenia gravis; rheumatoid and osteoarthritis; scleroderma; Chediak-Higashi and Sjogren's syndromes; systemic lupus erythematosus; toxic shock syndrome; traumatic tissue damage; and viral, bacterial, fungal, helminthic, and protozoal infections; a reproductive disorder such as disorders of prolactin production; infertility, including tubal disease, ovulatory defects, and endometriosis; disruptions of the estrous cycle, disruptions of the menstrual cycle, polycystic ovary syndrome, ovarian hyperstimulation syndrome, endometrial and ovarian tumors, uterine fibroids, autoimmune disorders, ectopic pregnancies, and teratogenesis; cancer of the breast, fibrocystic breast disease, and galactorrhea; disruptions of spermatogenesis, abnormal sperm physiology, cancer of the testis, cancer of the prostate, benign prostatic hyperplasia, prostatitis, Peyronie's disease, carcinoma of the male breast, and gynecomastia; a smooth muscle disorder such as any impairment or alteration in the normal action of smooth muscle including, but not limited to, that of the blood vessels, gastrointestinal tract, heart, and uterus, and including but not limited to, angina, anaphylactic shock, arrhythmias, asthma, cardiovascular shock, Cushing's syndrome, hypertension, hypoglycemia, myocardial infarction, migraine, and pheochromocytoma, and myopathies including cardiomyopathy, encephalopathy, epilepsy, Kearns-Sayre syndrome, lactic acidosis, myoclonic disorder, and ophthalmoplegia; a developmental disorder, such as any disorder associated with development or function of a tissue, organ, or system (such as the brain, adrenal gland, kidney, skeletal or reproductive system) of a subject, including but not limited to, renal tubular acidosis, anemia, Cushing's syndrome, achondroplastic dwarfism, Duchenne and Becker muscular dystrophy, epilepsy, gonadal dysgenesis, WAGR syndrome (Wilms' tumor, aniridia, genitourinary abnormalities, and mental retardation), Smith-Magenis syndrome, myelodysplastic syndrome, hereditary mucoepithelial dysplasia, hereditary keratodermas, hereditary neuropathies such as Charcot-Marie-Tooth disease and neurofibromatosis, hypothyroidism, hydrocephalus, seizure disorders such as Syndenham's chorea and cerebral palsy, spina bifida, anencephaly, craniorachischisis, congenital glaucoma, cataract, and sensorineural hearing loss; a nervous disorder such as akathesia, Alzheimer's disease, amnesia, amyotrophic lateral sclerosis and other motor neuron disorders, bipolar disorder, catatonia, cerebral neoplasms, dementia, depression, diabetic neuropathy, Down's syndrome, tardive dyskinesia, dystonias, epilepsy, Huntington's disease, peripheral neuropathy, multiple sclerosis, neurofibromatosis, Parkinson's disease and other extrapyramidal disorders, postherpetic neuralgia, epilepsy, ischemic cerebrovascular disease, stroke, cerebral neoplasms, Pick's disease, Huntington's disease, dementia, progressive neural muscular atrophy, retinitis pigmentosa, hereditary ataxias, multiple sclerosis and other demyelinating diseases, bacterial and viral meningitis, brain abscess, subdural empyema, epidural abscess, suppurative intracranial thrombophlebitis, myelitis and radiculitis, viral central nervous system disease; prion diseases including kuru, Creutzfeldt-Jakob disease, and Gerstmann-Straussler-Scheinker syndrome; fatal familial insomnia, nutritional and metabolic diseases of the nervous system, neurofibromatosis, tuberous sclerosis, cerebelloretinal

hemangioblastomatosis, encephalotrigeminal syndrome, mental retardation and other developmental disorders of the central nervous system, cerebral palsy, neuroskeletal disorders, autonomic nervous system disorders, cranial nerve disorders, spinal cord diseases, muscular dystrophy and other neuromuscular disorders, peripheral nervous system disorders, dermatomyositis and polymyositis; inherited, metabolic, endocrine, and toxic myopathies; myasthenia gravis, periodic paralysis; mental disorders including mood, anxiety, paranoid psychoses, and schizophrenic disorders; diabetic neuropathy, tardive dyskinesia, dystonias, and Tourette's disorder.

In another embodiment, a vector capable of expressing HCYT or a fragment or derivative thereof may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of HCYT including, but not limited to, those described above.

In a further embodiment, a pharmaceutical composition comprising a substantially purified HCYT in conjunction with a suitable pharmaceutical carrier may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of HCYT including, but not limited to, those provided above.

In still another embodiment, an agonist which modulates the activity of HCYT may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of HCYT including, but not limited to, those listed above.

In a further embodiment, an antagonist of HCYT may be administered to a subject to treat or prevent a disorder in which the expression or activity of HCYT is increased. Examples of such disorders include, but are not limited to, those cell proliferative, immunological, vesicle trafficking, reproductive, smooth muscle, developmental, and nervous disorders listed above. In one aspect, an antibody which specifically binds HCYT may be used directly as an antagonist or indirectly as a targeting or delivery mechanism for bringing a pharmaceutical agent to cells or tissue which express HCYT.

In an additional embodiment, a vector expressing the complement of the polynucleotide encoding HCYT may be administered to a subject to treat or prevent a disorder associated with increased expression or activity of HCYT including, but not limited to, those described above.

In other embodiments, any of the proteins, antagonists, antibodies, agonists, complementary sequences, or vectors of the invention may be administered in combination with other appropriate therapeutic agents. Selection of the appropriate agents for use in combination therapy may be made by one of ordinary skill in the art, according to conventional pharmaceutical principles. The combination of therapeutic agents may act synergistically to effect the treatment or prevention of the various disorders described above. Using this approach, one may be able to achieve therapeutic efficacy with lower dosages of each agent, thus reducing the potential for adverse side effects.

An antagonist of HCYT may be produced using methods which are generally known in the art. In particular, purified HCYT may be used to produce antibodies or to screen libraries of pharmaceutical agents to identify those which specifically bind HCYT. Antibodies to HCYT may also be generated using

methods that are well known in the art. Such antibodies may include, but are not limited to, polyclonal, monoclonal, chimeric, and single chain antibodies, Fab fragments, and fragments produced by a Fab expression library. Neutralizing antibodies (i.e., those which inhibit dimer formation) are especially preferred for therapeutic use.

For the production of antibodies, various hosts including goats, rabbits, rats, mice, humans, and others may be immunized by injection with HCYT or with any fragment or oligopeptide thereof which has immunogenic properties. Depending on the host species, various adjuvants may be used to increase immunological response. Such adjuvants include, but are not limited to, Freund's, mineral gels such as aluminum hydroxide, and surface active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, KLH, and dinitrophenol. Among adjuvants used in humans, BCG (bacilli Calmette-Guerin) and Corynebacterium parvum are especially preferable.

It is preferred that the oligopeptides, peptides, or fragments used to induce antibodies to HCYT have an amino acid sequence consisting of at least about 5 amino acids, and, more preferably, of at least about 10 amino acids. It is also preferable that these oligopeptides, peptides, or fragments are identical to a portion of the amino acid sequence of the natural protein and contain the entire amino acid sequence of a small, naturally occurring molecule. Short stretches of HCYT amino acids may be fused with those of another protein, such as KLH, and antibodies to the chimeric molecule may be produced.

Monoclonal antibodies to HCYT may be prepared using any technique which provides for the production of antibody molecules by continuous cell lines in culture. These include, but are not limited to, the hybridoma technique, the human B-cell hybridoma technique, and the EBV-hybridoma technique. (See, e.g., Kohler, G. et al. (1975) Nature 256:495-497; Kozbor, D. et al. (1985) J. Immunol. Methods 81:31-42; Cote, R.J. et al. (1983) Proc. Natl. Acad. Sci. 80:2026-2030; and Cole, S.P. et al. (1984) Mol. Cell Biol. 62:109-120.)

In addition, techniques developed for the production of "chimeric antibodies," such as the splicing of mouse antibody genes to human antibody genes to obtain a molecule with appropriate antigen specificity and biological activity, can be used. (See, e.g., Morrison, S.L. et al. (1984) Proc. Natl. Acad. Sci. 81:6851-6855; Neuberger, M.S. et al. (1984) Nature 312:604-608; and Takeda, S. et al. (1985) Nature 314:452-454.) Alternatively, techniques described for the production of single chain antibodies may be adapted, using methods known in the art, to produce HCYT-specific single chain antibodies. Antibodies with related specificity, but of distinct idiotypic composition, may be generated by chain shuffling from random combinatorial immunoglobulin libraries. (See, e.g., Burton D.R. (1991) Proc. Natl. Acad. Sci. 88:10134-10137.)

Antibodies may also be produced by inducing in vivo production in the lymphocyte population or by screening immunoglobulin libraries or panels of highly specific binding reagents as disclosed in the literature. (See, e.g., Orlandi, R. et al. (1989) Proc. Natl. Acad. Sci. 86: 3833-3837; Winter, G. et al. (1991) Nature 349:293-299.)

Antibody fragments which contain specific binding sites for HCYT may also be generated. For example, such fragments include, but are not limited to, F(ab')2 fragments produced by pepsin digestion of the antibody molecule and Fab fragments generated by reducing the disulfide bridges of the F(ab')2 fragments. Alternatively, Fab expression libraries may be constructed to allow rapid and easy identification of monoclonal Fab fragments with the desired specificity. (See, e.g., Huse, W.D. et al. (1989) Science 246:1275-1281.)

Various immunoassays may be used for screening to identify antibodies having the desired specificity. Numerous protocols for competitive binding or immunoradiometric assays using either polyclonal or monoclonal antibodies with established specificities are well known in the art. Such immunoassays typically involve the measurement of complex formation between HCYT and its specific antibody. A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering HCYT epitopes is preferred, but a competitive binding assay may also be employed (Pound, supra).

Various methods such as Scatchard analysis in conjunction with radioimmunoassay techniques may be used to assess the affinity of antibodies for HCYT. Affinity is expressed as an association constant, K_a, which is defined as the molar concentration of HCYT-antibody complex divided by the molar concentrations of free antigen and free antibody under equilibrium conditions. The K_a determined for a preparation of polyclonal antibodies, which are heterogeneous in their affinities for multiple HCYT epitopes, represents the average affinity, or avidity, of the antibodies for HCYT. The K_a determined for a preparation of monoclonal antibodies, which are monospecific for a particular HCYT epitope, represents a true measure of affinity. High-affinity antibody preparations with K_a ranging from about 10^9 to 10^{12} L/mole are preferred for use in immunoassays in which the HCYT-antibody complex must withstand rigorous manipulations. Low-affinity antibody preparations with K_a ranging from about 10^6 to 10^7 L/mole are preferred for use in immunopurification and similar procedures which ultimately require dissociation of HCYT, preferably in active form, from the antibody (Catty, D. (1988) Antibodies, Volume I: A Practical Approach, IRL Press, Washington DC; Liddell, J. E. and Cryer, A. (1991) A Practical Guide to Monoclonal Antibodies, John Wiley & Sons, New York NY).

The titer and avidity of polyclonal antibody preparations may be further evaluated to determine the quality and suitability of such preparations for certain downstream applications. For example, a polyclonal antibody preparation containing at least 1-2 mg specific antibody/ml, preferably 5-10 mg specific antibody/ml, is preferred for use in procedures requiring precipitation of HCYT-antibody complexes. Procedures for evaluating antibody specificity, titer, and avidity, and guidelines for antibody quality and usage in various applications, are generally available. (See, e.g., Catty, supra, and Coligan et al. supra.)

In another embodiment of the invention, the polynucleotides encoding HCYT, or any fragment or complement thereof, may be used for therapeutic purposes. In one aspect, the complement of the

polynucleotide encoding HCYT may be used in situations in which it would be desirable to block the transcription of the mRNA. In particular, cells may be transformed with sequences complementary to polynucleotides encoding HCYT. Thus, complementary molecules or fragments may be used to modulate HCYT activity, or to achieve regulation of gene function. Such technology is now well known in the art, and sense or antisense oligonucleotides or larger fragments can be designed from various locations along the coding or control regions of sequences encoding HCYT.

In another embodiment of the invention, the polynucleotides encoding HCYT, or any fragment or complement thereof, may be used for therapeutic purposes. In one aspect, the complement of the polynucleotide encoding HCYT may be used in situations in which it would be desirable to block the transcription of the mRNA. In particular, cells may be transformed with sequences complementary to polynucleotides encoding HCYT. Thus, complementary molecules or fragments may be used to modulate HCYT activity, or to achieve regulation of gene function. Such technology is now well known in the art, and sense or antisense oligonucleotides or larger fragments can be designed from various locations along the coding or control regions of sequences encoding HCYT.

Expression vectors derived from retroviruses, adenoviruses, or herpes or vaccinia viruses, or from various bacterial plasmids, may be used for delivery of nucleotide sequences to the targeted organ, tissue, or cell population. Methods which are well known to those skilled in the art can be used to construct vectors to express nucleic acid sequences complementary to the polynucleotides encoding HCYT. (See, e.g., Sambrook, supra; Ausubel, 1995, supra.)

Genes encoding HCYT can be turned off by transforming a cell or tissue with expression vectors which express high levels of a polynucleotide, or fragment thereof, encoding HCYT. Such constructs may be used to introduce untranslatable sense or antisense sequences into a cell. Even in the absence of integration into the DNA, such vectors may continue to transcribe RNA molecules until they are disabled by endogenous nucleases. Transient expression may last for a month or more with a non-replicating vector, and may last even longer if appropriate replication elements are part of the vector system.

As mentioned above, modifications of gene expression can be obtained by designing complementary sequences or antisense molecules (DNA, RNA, or PNA) to the control, 5', or regulatory regions of the gene encoding HCYT. Oligonucleotides derived from the transcription initiation site, e.g., between about positions -10 and +10 from the start site, are preferred. Similarly, inhibition can be achieved using triple helix base-pairing methodology. Triple helix pairing is useful because it causes inhibition of the ability of the double helix to open sufficiently for the binding of polymerases, transcription factors, or regulatory molecules. Recent therapeutic advances using triplex DNA have been described in the literature. (See, e.g., Gee, J.E. et al. (1994) in Huber, B.E. and B.I. Carr, Molecular and Immunologic Approaches, Futura Publishing, Mt. Kisco NY, pp. 163-177.) A complementary sequence or antisense molecule may also be designed to block translation of mRNA by preventing the transcript from binding to ribosomes.

Ribozymes, enzymatic RNA molecules, may also be used to catalyze the specific cleavage of RNA. The mechanism of ribozyme action involves sequence-specific hybridization of the ribozyme molecule to complementary target RNA, followed by endonucleolytic cleavage. For example, engineered hammerhead motif ribozyme molecules may specifically and efficiently catalyze endonucleolytic cleavage of sequences encoding HCYT.

Specific ribozyme cleavage sites within any potential RNA target are initially identified by scanning the target molecule for ribozyme cleavage sites, including the following sequences: GUA, GUU, and GUC. Once identified, short RNA sequences of between 15 and 20 ribonucleotides, corresponding to the region of the target gene containing the cleavage site, may be evaluated for secondary structural features which may render the oligonucleotide inoperable. The suitability of candidate targets may also be evaluated by testing accessibility to hybridization with complementary oligonucleotides using ribonuclease protection assays.

Complementary ribonucleic acid molecules and ribozymes of the invention may be prepared by any method known in the art for the synthesis of nucleic acid molecules. These include techniques for chemically synthesizing oligonucleotides such as solid phase phosphoramidite chemical synthesis. Alternatively, RNA molecules may be generated by in vitro and in vivo transcription of DNA sequences encoding HCYT. Such DNA sequences may be incorporated into a wide variety of vectors with suitable RNA polymerase promoters such as T7 or SP6. Alternatively, these cDNA constructs that synthesize complementary RNA, constitutively or inducibly, can be introduced into cell lines, cells, or tissues.

RNA molecules may be modified to increase intracellular stability and half-life. Possible modifications include, but are not limited to, the addition of flanking sequences at the 5' and/or 3' ends of the molecule, or the use of phosphorothioate or 2' O-methyl rather than phosphodiesterase linkages within the backbone of the molecule. This concept is inherent in the production of PNAs and can be extended in all of these molecules by the inclusion of nontraditional bases such as inosine, queosine, and wybutosine, as well as acetyl-, methyl-, thio-, and similarly modified forms of adenine, cytidine, guanine, thymine, and uridine which are not as easily recognized by endogenous endonucleases.

Many methods for introducing vectors into cells or tissues are available and equally suitable for use in vivo, in vitro, and ex vivo. For ex vivo therapy, vectors may be introduced into stem cells taken from the patient and clonally propagated for autologous transplant back into that same patient. Delivery by transfection, by liposome injections, or by polycationic amino polymers may be achieved using methods which are well known in the art. (See, e.g., Goldman, C.K. et al. (1997) Nature Biotechnology 15:462-466.)

Any of the therapeutic methods described above may be applied to any subject in need of such therapy, including, for example, mammals such as dogs, cats, cows, horses, rabbits, monkeys, and most preferably, humans.

An additional embodiment of the invention relates to the administration of a pharmaceutical or

sterile composition, in conjunction with a pharmaceutically acceptable carrier, for any of the therapeutic effects discussed above. Such pharmaceutical compositions may consist of HCYT, antibodies to HCYT, and mimetics, agonists, antagonists, or inhibitors of HCYT. The compositions may be administered alone or in combination with at least one other agent, such as a stabilizing compound, which may be administered in any sterile, biocompatible pharmaceutical carrier including, but not limited to, saline, buffered saline, dextrose, and water. The compositions may be administered to a patient alone, or in combination with other agents, drugs, or hormones.

The pharmaceutical compositions utilized in this invention may be administered by any number of routes including, but not limited to, oral, intravenous, intramuscular, intra-arterial, intramedullary, intrathecal, intraventricular, transdermal, subcutaneous, intraperitoneal, intranasal, enteral, topical, sublingual, or rectal means.

In addition to the active ingredients, these pharmaceutical compositions may contain suitable pharmaceutically-acceptable carriers comprising excipients and auxiliaries which facilitate processing of the active compounds into preparations which can be used pharmaceutically. Further details on techniques for formulation and administration may be found in the latest edition of Remington's Pharmaceutical Sciences (Maack Publishing, Easton PA).

Pharmaceutical compositions for oral administration can be formulated using pharmaceutically acceptable carriers well known in the art in dosages suitable for oral administration. Such carriers enable the pharmaceutical compositions to be formulated as tablets, pills, dragees, capsules, liquids, gels, syrups, slurries, suspensions, and the like, for ingestion by the patient.

Pharmaceutical preparations for oral use can be obtained through combining active compounds with solid excipient and processing the resultant mixture of granules (optionally, after grinding) to obtain tablets or dragee cores. Suitable auxiliaries can be added, if desired. Suitable excipients include carbohydrate or protein fillers, such as sugars, including lactose, sucrose, mannitol, and sorbitol; starch from corn, wheat, rice, potato, or other plants; cellulose, such as methyl cellulose, hydroxypropylmethyl-cellulose, or sodium carboxymethylcellulose; gums, including arabic and tragacanth; and proteins, such as gelatin and collagen. If desired, disintegrating or solubilizing agents may be added, such as the cross-linked polyvinyl pyrrolidone, agar, and alginic acid or a salt thereof, such as sodium alginate.

Dragee cores may be used in conjunction with suitable coatings, such as concentrated sugar solutions, which may also contain gum arabic, talc, polyvinylpyrrolidone, carbopol gel, polyethylene glycol, and/or titanium dioxide, lacquer solutions, and suitable organic solvents or solvent mixtures. Dyestuffs or pigments may be added to the tablets or dragee coatings for product identification or to characterize the quantity of active compound, i.e., dosage.

Pharmaceutical preparations which can be used orally include push-fit capsules made of gelatin, as well as soft, sealed capsules made of gelatin and a coating, such as glycerol or sorbitol. Push-fit

capsules can contain active ingredients mixed with fillers or binders, such as lactose or starches, lubricants, such as talc or magnesium stearate, and, optionally, stabilizers. In soft capsules, the active compounds may be dissolved or suspended in suitable liquids, such as fatty oils, liquid, or liquid polyethylene glycol with or without stabilizers.

Pharmaceutical formulations suitable for parenteral administration may be formulated in aqueous solutions, preferably in physiologically compatible buffers such as Hanks' solution, Ringer's solution, or physiologically buffered saline. Aqueous injection suspensions may contain substances which increase the viscosity of the suspension, such as sodium carboxymethyl cellulose, sorbitol, or dextran. Additionally, suspensions of the active compounds may be prepared as appropriate oily injection suspensions. Suitable lipophilic solvents or vehicles include fatty oils, such as sesame oil, or synthetic fatty acid esters, such as ethyl oleate, triglycerides, or liposomes. Non-lipid polycationic amino polymers may also be used for delivery. Optionally, the suspension may also contain suitable stabilizers or agents to increase the solubility of the compounds and allow for the preparation of highly concentrated solutions.

For topical or nasal administration, penetrants appropriate to the particular barrier to be permeated are used in the formulation. Such penetrants are generally known in the art.

The pharmaceutical compositions of the present invention may be manufactured in a manner that is known in the art, e.g., by means of conventional mixing, dissolving, granulating, dragee-making, levigating, emulsifying, encapsulating, entrapping, or lyophilizing processes.

The pharmaceutical composition may be provided as a salt and can be formed with many acids, including but not limited to, hydrochloric, sulfuric, acetic, lactic, tartaric, malic, and succinic acid. Salts tend to be more soluble in aqueous or other protonic solvents than are the corresponding free base forms. In other cases, the preferred preparation may be a lyophilized powder which may contain any or all of the following: 1 mM to 50 mM histidine, 0.1% to 2% sucrose, and 2% to 7% mannitol, at a pH range of 4.5 to 5.5, that is combined with buffer prior to use.

After pharmaceutical compositions have been prepared, they can be placed in an appropriate container and labeled for treatment of an indicated condition. For administration of HCYT, such labeling would include amount, frequency, and method of administration.

Pharmaceutical compositions suitable for use in the invention include compositions wherein the active ingredients are contained in an effective amount to achieve the intended purpose. The determination of an effective dose is well within the capability of those skilled in the art.

For any compound, the therapeutically effective dose can be estimated initially either in cell culture assays, e.g., of neoplastic cells or in animal models such as mice, rats, rabbits, dogs, or pigs. An animal model may also be used to determine the appropriate concentration range and route of administration. Such information can then be used to determine useful doses and routes for administration in humans.

A therapeutically effective dose refers to that amount of active ingredient, for example HCYT or

fragments thereof, antibodies of HCYT, and agonists, antagonists or inhibitors of HCYT, which ameliorates the symptoms or condition. Therapeutic efficacy and toxicity may be determined by standard pharmaceutical procedures in cell cultures or with experimental animals, such as by calculating the ED_{50} (the dose therapeutically effective in 50% of the population) or LD_{50} (the dose lethal to 50% of the population) statistics. The dose ratio of therapeutic to toxic effects is the therapeutic index, and it can be expressed as the ED_{50}/LD_{50} ratio. Pharmaceutical compositions which exhibit large therapeutic indices are preferred. The data obtained from cell culture assays and animal studies are used to formulate a range of dosage for human use. The dosage contained in such compositions is preferably within a range of circulating concentrations that includes the ED_{50} with little or no toxicity. The dosage varies within this range depending upon the dosage form employed, the sensitivity of the patient, and the route of administration.

The exact dosage will be determined by the practitioner, in light of factors related to the subject requiring treatment. Dosage and administration are adjusted to provide sufficient levels of the active moiety or to maintain the desired effect. Factors which may be taken into account include the severity of the disease state, the general health of the subject, the age, weight, and gender of the subject, time and frequency of administration, drug combination(s), reaction sensitivities, and response to therapy. Long-acting pharmaceutical compositions may be administered every 3 to 4 days, every week, or biweekly depending on the half-life and clearance rate of the particular formulation.

Normal dosage amounts may vary from about 0.1 μg to 100,000 μg, up to a total dose of about 1 gram, depending upon the route of administration. Guidance as to particular dosages and methods of delivery is provided in the literature and generally available to practitioners in the art. Those skilled in the art will employ different formulations for nucleotides than for proteins or their inhibitors. Similarly, delivery of polynucleotides or polypeptides will be specific to particular cells, conditions, locations, etc.

DIAGNOSTICS

In another embodiment, antibodies which specifically bind HCYT may be used for the diagnosis of cell proliferative, immunological, vesicle trafficking, reproductive, smooth muscle, developmental, and nervous disorders characterized by expression of HCYT, or in assays to monitor patients being treated with HCYT or agonists, antagonists, or inhibitors of HCYT. Antibodies useful for diagnostic purposes may be prepared in the same manner as described above for therapeutics. Diagnostic assays for HCYT include methods which utilize the antibody and a label to detect HCYT in human body fluids or in extracts of cells or tissues. The antibodies may be used with or without modification, and may be labeled by covalent or non-covalent attachment of a reporter molecule. A wide variety of reporter molecules, several of which are described above, are known in the art and may be used.

A variety of protocols for measuring HCYT, including ELISAs, RIAs, and FACS, are known in the art and provide a basis for diagnosing altered or abnormal levels of HCYT expression. Normal or standard values for HCYT expression are established by combining body fluids or cell extracts taken

from normal mammalian subjects, preferably human, with antibody to HCYT under conditions suitable for complex formation. The amount of standard complex formation may be quantitated by various methods, preferably by photometric means. Quantities of HCYT expressed in subject, control, and disease samples from biopsied tissues are compared with the standard values. Deviation between standard and subject values establishes the parameters for diagnosing disease.

In another embodiment of the invention, the polynucleotides encoding HCYT may be used for diagnostic purposes. The polynucleotides which may be used include oligonucleotide sequences, complementary RNA and DNA molecules, and PNAs. The polynucleotides may be used to detect and quantitate gene expression in biopsied tissues in which expression of HCYT may be correlated with disease. The diagnostic assay may be used to determine absence, presence, and excess expression of HCYT, and to monitor regulation of HCYT levels during therapeutic intervention.

In one aspect, hybridization with PCR probes which are capable of detecting polynucleotide sequences, including genomic sequences, encoding HCYT or closely related molecules may be used to identify nucleic acid sequences which encode HCYT. The specificity of the probe, whether it is made from a highly specific region, e.g., the 5' regulatory region, or from a less specific region, e.g., a conserved motif, and the stringency of the hybridization or amplification (maximal, high, intermediate, or low), will determine whether the probe identifies only naturally occurring sequences encoding HCYT, allelic variants, or related sequences.

Probes may also be used for the detection of related sequences, and should preferably have at least 50% sequence identity to any of the HCYT encoding sequences. The hybridization probes of the subject invention may be DNA or RNA and may be derived from the sequence of HCYT or from genomic sequences including promoters, enhancers, and introns of the HCYT gene.

Means for producing specific hybridization probes for DNAs encoding HCYT include the cloning of polynucleotide sequences encoding HCYT or HCYT derivatives into vectors for the production of mRNA probes. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes in vitro by means of the addition of the appropriate RNA polymerases and the appropriate labeled nucleotides. Hybridization probes may be labeled by a variety of reporter groups, for example, by radionuclides such as ^{32}P or ^{35}S, or by enzymatic labels, such as alkaline phosphatase coupled to the probe via avidin/biotin coupling systems, and the like.

Polynucleotide sequences encoding HCYT may be used for the diagnosis of cell proliferative, immunological, vesicle trafficking, reproductive, smooth muscle, developmental, and nervous disorders associated with expression of HCYT. Examples of such disorders include, but are not limited to, a cell proliferative disorder such as actinic keratosis, arteriosclerosis, atherosclerosis, bursitis, cirrhosis, hepatitis, mixed connective tissue disease (MCTD), myelofibrosis, paroxysmal nocturnal hemoglobinuria, polycythemia vera, psoriasis, primary thrombocythemia; cancers including adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in particular, cancers of the adrenal

gland, bladder, bone, bone marrow, brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus; an immunological disorder such as actinic keratosis, acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, arteriosclerosis, asthma, atherosclerosis, autoimmune hemolytic anemia, autoimmune thyroiditis, bronchitis, bursitis, cholecystitis, cirrhosis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, paroxysmal nocturnal hemoglobinuria, hepatitis, hypereosinophilia, irritable bowel syndrome, episodic lymphopenia with lymphocytotoxins, mixed connective tissue disease (MCTD), multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, myelofibrosis, osteoarthritis, osteoporosis, pancreatitis, polycythemia vera, polymyositis, psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome, systemic anaphylaxis, systemic lupus erythematosus, systemic sclerosis, primary thrombocythemia, thrombocytopenic purpura, ulcerative colitis, uveitis, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal circulation, viral, bacterial, fungal, parasitic, protozoal, and helminthic infections, and trauma; a vesicle trafficking disorder such as cystic fibrosis, glucose-galactose malabsorption syndrome, hypercholesterolemia, diabetes mellitus, diabetes insipidus, hyper- and hypoglycemia, Grave's disease, goiter, Cushing's disease, and Addison's disease; gastrointestinal disorders including ulcerative colitis, gastric and duodenal ulcers; other conditions associated with abnormal vesicle trafficking, including acquired immunodeficiency syndrome (AIDS); allergies including hay fever, asthma, and urticaria (hives); autoimmune hemolytic anemia; proliferative glomerulonephritis; inflammatory bowel disease; multiple sclerosis; myasthenia gravis; rheumatoid and osteoarthritis; scleroderma; Chediak-Higashi and Sjogren's syndromes; systemic lupus erythematosus; toxic shock syndrome; traumatic tissue damage; and viral, bacterial, fungal, helminthic, and protozoal infections; a reproductive disorder such as disorders of prolactin production; infertility, including tubal disease, ovulatory defects, and endometriosis; disruptions of the estrous cycle, disruptions of the menstrual cycle, polycystic ovary syndrome, ovarian hyperstimulation syndrome, endometrial and ovarian tumors, uterine fibroids, autoimmune disorders, ectopic pregnancies, and teratogenesis; cancer of the breast, fibrocystic breast disease, and galactorrhea; disruptions of spermatogenesis, abnormal sperm physiology, cancer of the testis, cancer of the prostate, benign prostatic hyperplasia, prostatitis, Peyronie's disease, carcinoma of the male breast, and gynecomastia; a smooth muscle disorder such as any impairment or alteration in the normal action of smooth muscle including, but not limited to, that of the blood vessels, gastrointestinal tract, heart, and uterus, and including but not limited to, angina, anaphylactic shock, arrhythmias, asthma, cardiovascular shock, Cushing's syndrome, hypertension, hypoglycemia, myocardial infarction, migraine, and pheochromocytoma, and myopathies including cardiomyopathy, encephalopathy, epilepsy, Kearns-Sayre

syndrome, lactic acidosis, myoclonic disorder, and ophthalmoplegia; a developmental disorder, such as any disorder associated with development or function of a tissue, organ, or system (such as the brain, adrenal gland, kidney, skeletal or reproductive system) of a subject, including but not limited to, renal tubular acidosis, anemia, Cushing's syndrome, achondroplastic dwarfism, Duchenne and Becker muscular dystrophy, epilepsy, gonadal dysgenesis, WAGR syndrome (Wilms' tumor, aniridia, genitourinary abnormalities, and mental retardation), Smith-Magenis syndrome, myelodysplastic syndrome, hereditary mucoepithelial dysplasia, hereditary keratodermas, hereditary neuropathies such as Charcot-Marie-Tooth disease and neurofibromatosis, hypothyroidism, hydrocephalus, seizure disorders such as Syndenham's chorea and cerebral palsy, spina bifida, anencephaly, craniorachischisis, congenital glaucoma, cataract, and sensorineural hearing loss; a nervous disorder such as akathesia, Alzheimer's disease, amnesia, amyotrophic lateral sclerosis and other motor neuron disorders, bipolar disorder, catatonia, cerebral neoplasms, dementia, depression, diabetic neuropathy, Down's syndrome, tardive dyskinesia, dystonias, epilepsy, Huntington's disease, peripheral neuropathy, multiple sclerosis, neurofibromatosis, Parkinson's disease and other extrapyramidal disorders, postherpetic neuralgia, epilepsy, ischemic cerebrovascular disease, stroke, cerebral neoplasms, Pick's disease, Huntington's disease, dementia, progressive neural muscular atrophy, retinitis pigmentosa, hereditary ataxias, multiple sclerosis and other demyelinating diseases, bacterial and viral meningitis, brain abscess, subdural empyema, epidural abscess, suppurative intracranial thrombophlebitis, myelitis and radiculitis, viral central nervous system disease; prion diseases including kuru, Creutzfeldt-Jakob disease, and Gerstmann-Straussler-Scheinker syndrome; fatal familial insomnia, nutritional and metabolic diseases of the nervous system, neurofibromatosis, tuberous sclerosis, cerebelloretinal hemangioblastomatosis, encephalotrigeminal syndrome, mental retardation and other developmental disorders of the central nervous system, cerebral palsy, neuroskeletal disorders, autonomic nervous system disorders, cranial nerve disorders, spinal cord diseases, muscular dystrophy and other neuromuscular disorders, peripheral nervous system disorders, dermatomyositis and polymyositis; inherited, metabolic, endocrine, and toxic myopathies; myasthenia gravis, periodic paralysis; mental disorders including mood, anxiety, paranoid psychoses, and schizophrenic disorders; diabetic neuropathy, tardive dyskinesia, dystonias, and Tourette's disorder. The polynucleotide sequences encoding HCYT may be used in Southern or northern analysis, dot blot, or other membrane-based technologies; in PCR technologies; in dipstick, pin, and ELISA assays; and in microarrays utilizing fluids or tissues from patients to detect altered HCYT expression. Such qualitative or quantitative methods are well known in the art.

In a particular aspect, the nucleotide sequences encoding HCYT may be useful in assays that detect the presence of associated disorders, particularly those mentioned above. The nucleotide sequences encoding HCYT may be labeled by standard methods and added to a fluid or tissue sample from a patient under conditions suitable for the formation of hybridization complexes. After a suitable incubation period, the sample is washed and the signal is quantitated and compared with a standard value.

If the amount of signal in the patient sample is significantly altered in comparison to a control sample then the presence of altered levels of nucleotide sequences encoding HCYT in the sample indicates the presence of the associated disorder. Such assays may also be used to evaluate the efficacy of a particular therapeutic treatment regimen in animal studies, in clinical trials, or to monitor the treatment of an individual patient.

In order to provide a basis for the diagnosis of a disorder associated with expression of HCYT, a normal or standard profile for expression is established. This may be accomplished by combining body fluids or cell extracts taken from normal subjects, either animal or human, with a sequence, or a fragment thereof, encoding HCYT, under conditions suitable for hybridization or amplification. Standard hybridization may be quantified by comparing the values obtained from normal subjects with values from an experiment in which a known amount of a substantially purified polynucleotide is used. Standard values obtained in this manner may be compared with values obtained from samples from patients who are symptomatic for a disorder. Deviation from standard values is used to establish the presence of a disorder.

Once the presence of a disorder is established and a treatment protocol is initiated, hybridization assays may be repeated on a regular basis to determine if the level of expression in the patient begins to approximate that which is observed in the normal subject. The results obtained from successive assays may be used to show the efficacy of treatment over a period ranging from several days to months.

With respect to cancer, the presence of an abnormal amount of transcript (either under- or overexpressed) in biopsied tissue from an individual may indicate a predisposition for the development of the disease, or may provide a means for detecting the disease prior to the appearance of actual clinical symptoms. A more definitive diagnosis of this type may allow health professionals to employ preventative measures or aggressive treatment earlier thereby preventing the development or further progression of the cancer.

Additional diagnostic uses for oligonucleotides designed from the sequences encoding HCYT may involve the use of PCR. These oligomers may be chemically synthesized, generated enzymatically, or produced in vitro. Oligomers will preferably contain a fragment of a polynucleotide encoding HCYT, or a fragment of a polynucleotide complementary to the polynucleotide encoding HCYT, and will be employed under optimized conditions for identification of a specific gene or condition. Oligomers may also be employed under less stringent conditions for detection or quantitation of closely related DNA or RNA sequences.

Methods which may also be used to quantitate the expression of HCYT include radiolabeling or biotinylating nucleotides, coamplification of a control nucleic acid, and interpolating results from standard curves. (See, e.g., Melby, P.C. et al. (1993) J. Immunol. Methods 159:235-244; Duplaa, C. et al. (1993) Anal. Biochem. 229-236.) The speed of quantitation of multiple samples may be accelerated by running the assay in an ELISA format where the oligomer of interest is presented in various dilutions and a

spectrophotometric or colorimetric response gives rapid quantitation.

In further embodiments, oligonucleotides or longer fragments derived from any of the polynucleotide sequences described herein may be used as targets in a microarray. The microarray can be used to monitor the expression level of large numbers of genes simultaneously and to identify genetic variants, mutations, and polymorphisms. This information may be used to determine gene function, to understand the genetic basis of a disorder, to diagnose a disorder, and to develop and monitor the activities of therapeutic agents.

Microarrays may be prepared, used, and analyzed using methods known in the art. (See, e.g., Brennan, T.M. et al. (1995) U.S. Patent No. 5,474,796; Schena, M. et al. (1996) Proc. Natl. Acad. Sci. 93:10614-10619; Baldeschweiler et al. (1995) PCT application WO95/251116; Shalon, D. et al. (1995) PCT application WO95/35505; Heller, R.A. et al. (1997) Proc. Natl. Acad. Sci. 94:2150-2155; and Heller, M.J. et al. (1997) U.S. Patent No. 5,605,662.)

In another embodiment of the invention, nucleic acid sequences encoding HCYT may be used to generate hybridization probes useful in mapping the naturally occurring genomic sequence. The sequences may be mapped to a particular chromosome, to a specific region of a chromosome, or to artificial chromosome constructions, e.g., human artificial chromosomes (HACs), yeast artificial chromosomes (YACs), bacterial artificial chromosomes (BACs), bacterial P1 constructions, or single chromosome cDNA libraries. (See, e.g., Harrington, J.J. et al. (1997) Nat Genet. 15:345-355; Price, C.M. (1993) Blood Rev. 7:127-134; and Trask, B.J. (1991) Trends Genet. 7:149-154.)

Fluorescent in situ hybridization (FISH) may be correlated with other physical chromosome mapping techniques and genetic map data. (See, e.g., Heinz-Ulrich, et al. (1995) in Meyers, supra, pp. 965-968.) Examples of genetic map data can be found in various scientific journals or at the Online Mendelian Inheritance in Man (OMIM) site. Correlation between the location of the gene encoding HCYT on a physical chromosomal map and a specific disorder, or a predisposition to a specific disorder, may help define the region of DNA associated with that disorder. The nucleotide sequences of the invention may be used to detect differences in gene sequences among normal, carrier, and affected individuals.

In situ hybridization of chromosomal preparations and physical mapping techniques, such as linkage analysis using established chromosomal markers, may be used for extending genetic maps. Often the placement of a gene on the chromosome of another mammalian species, such as mouse, may reveal associated markers even if the number or arm of a particular human chromosome is not known. New sequences can be assigned to chromosomal arms by physical mapping. This provides valuable information to investigators searching for disease genes using positional cloning or other gene discovery techniques. Once the disease or syndrome has been crudely localized by genetic linkage to a particular genomic region, e.g., ataxia-telangiectasia to 11q22-23, any sequences mapping to that area may represent associated or regulatory genes for further investigation. (See, e.g., Gatti, R.A. et al. (1988) Nature

336:577-580.) The nucleotide sequence of the subject invention may also be used to detect differences in the chromosomal location due to translocation, inversion, etc., among normal, carrier, or affected individuals.

In another embodiment of the invention, HCYT, its catalytic or immunogenic fragments, or oligopeptides thereof can be used for screening libraries of compounds in any of a variety of drug screening techniques. The fragment employed in such screening may be free in solution, affixed to a solid support, borne on a cell surface, or located intracellularly. The formation of binding complexes between HCYT and the agent being tested may be measured.

Another technique for drug screening provides for high throughput screening of compounds having suitable binding affinity to the protein of interest. (See, e.g., Geysen, et al. (1984) PCT application WO84/03564.) In this method, large numbers of different small test compounds are synthesized on a solid substrate. The test compounds are reacted with HCYT, or fragments thereof, and washed. Bound HCYT is then detected by methods well known in the art. Purified HCYT can also be coated directly onto plates for use in the aforementioned drug screening techniques. Alternatively, non-neutralizing antibodies can be used to capture the peptide and immobilize it on a solid support.

In another embodiment, one may use competitive drug screening assays in which neutralizing antibodies capable of binding HCYT specifically compete with a test compound for binding HCYT. In this manner, antibodies can be used to detect the presence of any peptide which shares one or more antigenic determinants with HCYT.

In additional embodiments, the nucleotide sequences which encode HCYT may be used in any molecular biology techniques that have yet to be developed, provided the new techniques rely on properties of nucleotide sequences that are currently known, including, but not limited to, such properties as the triplet genetic code and specific base pair interactions.

Without further elaboration, it is believed that one skilled in the art can, using the preceding description, utilize the present invention to its fullest extent. The following preferred specific embodiments are, therefore, to be construed as merely illustrative, and not limitative of the remainder of the disclosure in any way whatsoever.

The disclosures of all patents, applications, and publications mentioned above and below, in particular U.S. Ser. No. [Atty Docket No. PF-0566 P, filed August 4, 1998], U.S. Ser. No. [Atty Docket No. PF-0568 P, filed July 31, 1998] and U.S. Ser. No.: [Atty Docket No. PF-0578 P, filed August 19, 1998], are hereby expressly incorporated by reference.

EXAMPLES

I. Construction of cDNA Libraries

RNA was purchased from Clontech or isolated from tissues described in Table 4. Some tissues were homogenized and lysed in guanidinium isothiocyanate, while others were homogenized and lysed in

phenol or in a suitable mixture of denaturants, such as TRIZOL (Life Technologies), a monophasic solution of phenol and guanidine isothiocyanate. The resulting lysates were centrifuged over CsCl cushions or extracted with chloroform. RNA was precipitated from the lysates with either isopropanol or sodium acetate and ethanol, or by other routine methods.

Phenol extraction and precipitation of RNA were repeated as necessary to increase RNA purity. In some cases, RNA was treated with DNase. For most libraries, poly(A+) RNA was isolated using oligo d(T)-coupled paramagnetic particles (Promega), OLIGOTEX latex particles (QIAGEN, Valencia CA), or an OLIGOTEX mRNA purification kit (QIAGEN). Alternatively, RNA was isolated directly from tissue lysates using other RNA isolation kits, e.g., the POLY(A)PURE mRNA purification kit (Ambion, Austin TX).

In some cases, Stratagene was provided with RNA and constructed the corresponding cDNA libraries. Otherwise, cDNA was synthesized and cDNA libraries were constructed with the UNIZAP vector system (Stratagene) or SUPERSCRIPT plasmid system (Life Technologies), using the recommended procedures or similar methods known in the art. (See, e.g., Ausubel, 1997, supra, units 5.1-6.6). Reverse transcription was initiated using oligo d(T) or random primers. Synthetic oligonucleotide adapters were ligated to double stranded cDNA, and the cDNA was digested with the appropriate restriction enzyme or enzymes. For most libraries, the cDNA was size-selected (300-1000 bp) using SEPHACRYL S1000, SEPHAROSE CL2B, or SEPHAROSE CL4B column chromatography (Amersham Pharmacia Biotech) or preparative agarose gel electrophoresis. cDNAs were ligated into compatible restriction enzyme sites of the polylinker of a suitable plasmid, e.g., PBLUESCRIPT plasmid (Stratagene), pSPORT1 plasmid (Life Technologies), or pINCY (Incyte Pharmaceuticals, Palo Alto CA). Recombinant plasmids were transformed into competent E. coli cells including XL1-Blue, XL1-BlueMRF, or SOLR from Stratagene or DH5α, DH10B, or ElectroMAX DH10B from Life Technologies.

II. Isolation of cDNA Clones

Plasmids were recovered from host cells by in vivo excision, using the UNIZAP vector system (Stratagene) or cell lysis. Plasmids were purified using at least one of the following: a Magic or WIZARD Minipreps DNA purification system (Promega); an AGTC Miniprep purification kit (Edge Biosystems, Gaithersburg MD); and QIAWELL 8 Plasmid, QIAWELL 8 Plus Plasmid, QIAWELL 8 Ultra Plasmid purification systems or the REAL Prep 96 plasmid kit from QIAGEN. Following precipitation, plasmids were resuspended in 0.1 ml of distilled water and stored, with or without lyophilization, at 4°C.

Alternatively, plasmid DNA was amplified from host cell lysates using direct link PCR in a high-throughput format (Rao, V.B. (1994) Anal. Biochem. 216:1-14). Host cell lysis and thermal cycling steps were carried out in a single reaction mixture. Samples were processed and stored in 384-well plates, and the concentration of amplified plasmid DNA was quantified fluorometrically using PICOGREEN reagent (Molecular Probes, Eugene OR) and a Fluoroskan II fluorescence scanner (Labsystems Oy, Helsinki,

Finland).

III. Sequencing and Analysis

The cDNAs were prepared for sequencing using the ABI CATALYST 800 (Perkin-Elmer) or the HYDRA microdispenser (Robbins Scientific) or MICROLAB 2200 (Hamilton) systems in combination with the PTC-200 thermal cyclers (MJ Research). The cDNAs were sequenced using the ABI PRISM 373 or 377 sequencing systems (Perkin-Elmer) and standard ABI protocols, base calling software, and kits. In one alternative, cDNAs were sequenced using the MEGABACE 1000 DNA sequencing system (Molecular Dynamics). In another alternative, the cDNAs were amplified and sequenced using the ABI PRISM BIGDYE Terminator cycle sequencing ready reaction kit (Perkin-Elmer). In yet another alternative, cDNAs were sequenced using solutions and dyes from Amersham Pharmacia Biotech. Reading frames for the ESTs were determined using standard methods as reviewed in Ausubel (1997, supra, unit 7.7). Some of the cDNA sequences were selected for extension using the techniques disclosed in Example V.

The polynucleotide sequences derived from cDNA, extension, and shotgun sequencing were assembled and analyzed using a combination of software programs which utilize algorithms well known to those skilled in the art. Table 5 summarizes the software programs, descriptions, references, and threshold parameters used. The first column of Table 5 shows the tools, programs, and algorithms used, the second column provides a brief description thereof, the third column presents the references which are incorporated by reference herein, and the fourth column presents, where applicable, the scores, probability values, and other parameters used to evaluate the strength of a match between two sequences (the higher the probability the greater the homology). Sequences were analyzed using MACDNASIS PRO software (Hitachi Software Engineering, S. San Francisco CA) and LASERGENE software (DNASTAR).

The polynucleotide sequences were validated by removing vector, linker, and polyA sequences and by masking ambiguous bases, using algorithms and programs based on BLAST, dynamic programing, and dinucleotide nearest neighbor analysis. The sequences were then queried against a selection of public databases such as GenBank primate, rodent, mammalian, vertebrate, and eukaryote databases, and BLOCKS to acquire annotation, using programs based on BLAST, FASTA, and BLIMPS. The sequences were assembled into full length polynucleotide sequences using programs based on Phred, Phrap, and Consed, and were screened for open reading frames using programs based on GeneMark, BLAST, and FASTA. The full length polynucleotide sequences were translated to derive the corresponding full length amino acid sequences, and these full length sequences were subsequently analyzed by querying against databases such as the GenBank databases (described above), SwissProt, BLOCKS, PRINTS, PFAM, and Prosite.

IV. Northern Analysis

Northern analysis is a laboratory technique used to detect the presence of a transcript of a gene and involves the hybridization of a labeled nucleotide sequence to a membrane on which RNAs from a particular cell type or tissue have been bound. (See, e.g., Sambrook, supra, ch. 7; Ausubel, 1995, supra, ch. 4 and 16.)

Analogous computer techniques applying BLAST were used to search for identical or related molecules in nucleotide databases such as GenBank or LIFESEQ database (Incyte Pharmaceuticals). This analysis is much faster than multiple membrane-based hybridizations. In addition, the sensitivity of the computer search can be modified to determine whether any particular match is categorized as exact or similar. The basis of the search is the product score, which is defined as:

$$\frac{\%\ \text{sequence identity} \times \%\ \text{maximum BLAST score}}{100}$$

The product score takes into account both the degree of similarity between two sequences and the length of the sequence match. For example, with a product score of 40, the match will be exact within a 1% to 2% error, and, with a product score of 70, the match will be exact. Similar molecules are usually identified by selecting those which show product scores between 15 and 40, although lower scores may identify related molecules.

The results of northern analyses are reported a percentage distribution of libraries in which the transcript encoding HCYT occurred. Analysis involved the categorization of cDNA libraries by organ/tissue and disease. The organ/tissue categories included cardiovascular, dermatologic, developmental, endocrine, gastrointestinal, hematopoietic/immune, musculoskeletal, nervous, reproductive, and urologic. The disease categories included cancer, inflammation/trauma, fetal, neurological, and pooled. For each category, the number of libraries expressing the sequence of interest was counted and divided by the total number of libraries across all categories. Percentage values of tissue-specific and disease expression are reported in Table 3.

V. Extension of HCYT Encoding Polynucleotides

Full length nucleic acid sequences (SEQ ID NO:9-14) were produced by extension of the component fragments described in Table 1, Column 5, using oligonucleotide primers based on those fragments. Primers were used to facilitate the extension of the known sequence "outward" generating amplicons containing new unknown nucleotide sequence for the region of interest. The initial primers were designed from the cDNA using OLIGO 4.06 (National Biosciences, Inc.), or another appropriate program, to be about 22 to 30 nucleotides in length, to have a GC content of about 50% or more, and to anneal to the target sequence at temperatures of about 68°C to about 72°C. Any stretch of nucleotides which would result in hairpin structures and primer-primer dimerizations was avoided.

Selected human cDNA libraries were used to extend the sequence. If more than one extension is necessary or desired, additional sets of primers are designed to further extend the known region.

High fidelity amplification was obtained by following the instructions for the XL-PCR kit (The

Perkin-Elmer Corp.) and thoroughly mixing the enzyme and reaction mix. PCR was performed using the PTC-200 thermal cycler (MJ Research, Inc.), beginning with 40 pmol of each primer and the recommended concentrations of all other components of the kit, with the following parameters:

Step 1	94° C for 1 min (initial denaturation)
Step 2	65° C for 1 min
Step 3	68° C for 6 min
Step 4	94° C for 15 sec
Step 5	65° C for 1 min
Step 6	68° C for 7 min
Step 7	Repeat steps 4 through 6 for an additional 15 cycles
Step 8	94° C for 15 sec
Step 9	65° C for 1 min
Step 10	68° C for 7:15 min
Step 11	Repeat steps 8 through 10 for an additional 12 cycles
Step 12	72° C for 8 min
Step 13	4° C (and holding)

A 5 μl to 10 μl aliquot of the reaction mixture was analyzed by electrophoresis on a low concentration (about 0.6% to 0.8%) agarose mini-gel to determine which reactions were successful in extending the sequence. Bands thought to contain the largest products were excised from the gel, purified using QIAQUICK™ (QIAGEN Inc.), and trimmed of overhangs using Klenow enzyme to facilitate religation and cloning.

After ethanol precipitation, the products were redissolved in 13 μl of ligation buffer, 1μl T4-DNA ligase (15 units) and 1μl T4 polynucleotide kinase were added, and the mixture was incubated at room temperature for 2 to 3 hours, or overnight at 16° C. Competent E. coli cells (in 40 μl of appropriate media) were transformed with 3 μl of ligation mixture and cultured in 80 μl of SOC medium. (See, e.g., Sambrook, supra, Appendix A, p. 2.) After incubation for one hour at 37°C, the E. coli mixture was plated on Luria Bertani (LB) agar (See, e.g., Sambrook, supra, Appendix A, p. 1) containing carbenicillin (2x carb). The following day, several colonies were randomly picked from each plate and cultured in 150 μl of liquid LB/2x carb medium placed in an individual well of an appropriate commercially-available sterile 96-well microtiter plate. The following day, 5 μl of each overnight culture was transferred into a non-sterile 96-well plate and, after dilution 1:10 with water, 5 μl from each sample was transferred into a PCR array.

For PCR amplification, 18 μl of concentrated PCR reaction mix (3.3x) containing 4 units of rTth DNA polymerase, a vector primer, and one or both of the gene specific primers used for the extension reaction were added to each well. Amplification was performed using the following conditions:

Step 1	94° C for 60 sec
Step 2	94° C for 20 sec
Step 3	55° C for 30 sec
Step 4	72° C for 90 sec
Step 5	Repeat steps 2 through 4 for an additional 29 cycles
Step 6	72° C for 180 sec
Step 7	4° C (and holding)

Aliquots of the PCR reactions were run on agarose gels together with molecular weight markers. The sizes of the PCR products were compared to the original partial cDNAs, and appropriate clones were selected, ligated into plasmid, and sequenced.

In like manner, the nucleotide sequence of SEQ ID NO:9-14 are used to obtain 5' regulatory sequences using the procedure above, oligonucleotides designed for 5' extension, and an appropriate genomic library.

The full length nucleic acid sequences of SEQ ID NO:15-16 were produced by extension of an appropriate fragment of the full length molecule using oligonucleotide primers designed from this fragment. One primer was synthesized to initiate 5' extension of the known fragment, and the other primer, to initiate 3' extension of the known fragment. The initial primers were designed using OLIGO 4.06 software (National Biosciences), or another appropriate program, to be about 22 to 30 nucleotides in length, to have a GC content of about 50% or more, and to anneal to the target sequence at temperatures of about 68°C to about 72°C. Any stretch of nucleotides which would result in hairpin structures and primer-primer dimerizations was avoided.

Selected human cDNA libraries were used to extend the sequence. If more than one extension was necessary or desired, additional or nested sets of primers were designed.

High fidelity amplification was obtained by PCR using methods well known in the art. PCR was performed in 96-well plates using the PTC-200 thermal cycler (MJ Research, Inc.). The reaction mix contained DNA template, 200 nmol of each primer, reaction buffer containing Mg^{2+}, $(NH_4)_2SO_4$, and β-mercaptoethanol, Taq DNA polymerase (Amersham Pharmacia Biotech), ELONGASE enzyme (Life Technologies), and Pfu DNA polymerase (Stratagene), with the following parameters for primer pair PCI A and PCI B: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 60°C, 1 min; Step 4: 68°C, 2 min; Step 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68°C, 5 min; Step 7: storage at 4°C. In the alternative, the parameters for primer pair T7 and SK+ were as follows: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 57°C, 1 min; Step 4: 68°C, 2 min; Step 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68°C, 5 min; Step 7: storage at 4°C.

The concentration of DNA in each well was determined by dispensing 100 µl PICOGREEN quantitation reagent (0.25% (v/v) PICOGREEN; Molecular Probes) dissolved in 1X TE and 0.5 µl of undiluted PCR product into each well of an opaque fluorimeter plate (Corning Costar, Acton MA), allowing the DNA to bind to the reagent. The plate was scanned in a Fluoroskan II (Labsystems Oy, Helsinki, Finland) to measure the fluorescence of the sample and to quantify the concentration of DNA. A 5 µl to 10 µl aliquot of the reaction mixture was analyzed by electrophoresis on a 1 % agarose mini-gel to determine which reactions were successful in extending the sequence.

The extended nucleotides were desalted and concentrated, transferred to 384-well plates, digested with CviJI cholera virus endonuclease (Molecular Biology Research, Madison WI), and sonicated or sheared prior to religation into pUC 18 vector (Amersham Pharmacia Biotech). For shotgun

sequencing, the digested nucleotides were separated on low concentration (0.6 to 0.8%) agarose gels, fragments were excised, and agar digested with Agar ACE (Promega). Extended clones were religated using T4 ligase (New England Biolabs, Beverly MA) into pUC 18 vector (Amersham Pharmacia Biotech), treated with Pfu DNA polymerase (Stratagene) to fill-in restriction site overhangs, and transfected into competent E. coli cells. Transformed cells were selected on antibiotic-containing media, individual colonies were picked and cultured overnight at 37°C in 384-well plates in LB/2x carb liquid media.

The cells were lysed, and DNA was amplified by PCR using Taq DNA polymerase (Amersham Pharmacia Biotech) and Pfu DNA polymerase (Stratagene) with the following parameters: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 60°C, 1 min; Step 4: 72°C, 2 min; Step 5: steps 2, 3, and 4 repeated 29 times; Step 6: 72°C, 5 min; Step 7: storage at 4°C. DNA was quantified by PICOGREEN reagent (Molecular Probes) as described above. Samples with low DNA recoveries were reamplified using the same conditions as described above. Samples were diluted with 20% dimethysulphoxide (1:2, v/v), and sequenced using DYENAMIC energy transfer sequencing primers and the DYENAMIC DIRECT kit (Amersham Pharmacia Biotech) or the ABI PRISM BIGDYE Terminator cycle sequencing ready reaction kit (Perkin-Elmer).

In like manner, the nucleotide sequences of SEQ ID NO:15-16 are used to obtain 5' regulatory sequences using the procedure above, oligonucleotides designed for such extension, and an appropriate genomic library.

VI. Labeling and Use of Individual Hybridization Probes

Hybridization probes derived from SEQ ID NO:9-16 are employed to screen cDNAs, genomic DNAs, or mRNAs. Although the labeling of oligonucleotides, consisting of about 20 base pairs, is specifically described, essentially the same procedure is used with larger nucleotide fragments. Oligonucleotides are designed using state-of-the-art software such as OLIGO 4.06 software (National Biosciences, Inc.) and labeled by combining 50 pmol of each oligomer, 250 μCi of [γ-^{32}P] adenosine triphosphate (Amersham Pharmacia Biotech, Ltd.), and T4 polynucleotide kinase (DuPont NEN, Boston, MA). The labeled oligonucleotides are substantially purified using a SEPHADEX G-25 superfine size exclusion dextran bead column (Amersham Pharmacia Biotech, Ltd.). An aliquot containing 10^7 counts per minute of the labeled probe is used in a typical membrane-based hybridization analysis of human genomic DNA digested with one of the following endonucleases: Ase I, Bgl II, Eco RI, Pst I, Xba 1, or Pvu II (DuPont NEN, Boston, MA).

The DNA from each digest is fractionated on a 0.7% agarose gel and transferred to nylon membranes (NYTRAN Plus, Schleicher & Schuell, Durham, NH). Hybridization is carried out for 16 hours at 40°C. To remove nonspecific signals, blots are sequentially washed at room temperature under increasingly stringent conditions up to 0.1 x saline sodium citrate and 0.5% sodium dodecyl sulfate. After XOMAT AR film (Eastman Kodak, Rochester, NY) is exposed to the blots, hybridization patterns are

compared visually.

VII. Microarrays

A chemical coupling procedure and an ink jet device can be used to synthesize array elements on the surface of a substrate. (See, e.g., Baldeschweiler, supra.) An array analogous to a dot or slot blot may also be used to arrange and link elements to the surface of a substrate using thermal, UV, chemical, or mechanical bonding procedures. A typical array may be produced by hand or using available methods and machines and contain any appropriate number of elements. After hybridization, nonhybridized probes are removed and a scanner used to determine the levels and patterns of fluorescence. The degree of complementarity and the relative abundance of each probe which hybridizes to an element on the microarray may be assessed through analysis of the scanned images.

Full-length cDNAs, Expressed Sequence Tags (ESTs), or fragments thereof may comprise the elements of the microarray. Fragments suitable for hybridization can be selected using software well known in the art such as LASERGENE software (DNASTAR). Full-length cDNAs, ESTs, or fragments thereof corresponding to one of the nucleotide sequences of the present invention, or selected at random from a cDNA library relevant to the present invention, are arranged on an appropriate substrate, e.g., a glass slide. The cDNA is fixed to the slide using, e.g., UV cross-linking followed by thermal and chemical treatments and subsequent drying. (See, e.g., Schena, M. et al. (1995) Science 270:467-470; Shalon, D. et al. (1996) Genome Res. 6:639-645.) Fluorescent probes are prepared and used for hybridization to the elements on the substrate. The substrate is analyzed by procedures described above.

VIII. Complementary Polynucleotides

Sequences complementary to the HCYT-encoding sequences, or any parts thereof, are used to detect, decrease, or inhibit expression of naturally occurring HCYT. Although use of oligonucleotides comprising from about 15 to 30 base pairs is described, essentially the same procedure is used with smaller or with larger sequence fragments. Appropriate oligonucleotides are designed using OLIGO 4.06 software (National Biosciences) and the coding sequence of HCYT. To inhibit transcription, a complementary oligonucleotide is designed from the most unique 5' sequence and used to prevent promoter binding to the coding sequence. To inhibit translation, a complementary oligonucleotide is designed to prevent ribosomal binding to the HCYT-encoding transcript.

IX. Expression of HCYT

Expression and purification of HCYT is achieved using bacterial or virus-based expression systems. For expression of HCYT in bacteria, cDNA is subcloned into an appropriate vector containing an antibiotic resistance gene and an inducible promoter that directs high levels of cDNA transcription. Examples of such promoters include, but are not limited to, the *trp-lac* (*tac*) hybrid promoter and the T5 or T7 bacteriophage promoter in conjunction with the *lac* operator regulatory element. Recombinant vectors are transformed into suitable bacterial hosts, e.g., BL21(DE3). Antibiotic resistant bacteria express HCYT upon induction with isopropyl beta-D-thiogalactopyranoside (IPTG). Expression of

HCYT in eukaryotic cells is achieved by infecting insect or mammalian cell lines with recombinant Autographica californica nuclear polyhedrosis virus (AcMNPV), commonly known as baculovirus. The nonessential polyhedrin gene of baculovirus is replaced with cDNA encoding HCYT by either homologous recombination or bacterial-mediated transposition involving transfer plasmid intermediates. Viral infectivity is maintained and the strong polyhedrin promoter drives high levels of cDNA transcription. Recombinant baculovirus is used to infect Spodoptera frugiperda (Sf9) insect cells in most cases, or human hepatocytes, in some cases. Infection of the latter requires additional genetic modifications to baculovirus. (See Engelhard, E. K. et al. (1994) Proc. Natl. Acad. Sci. USA 91:3224-3227; Sandig, V. et al. (1996) Hum. Gene Ther. 7:1937-1945.)

In most expression systems, HCYT is synthesized as a fusion protein with, e.g., glutathione S-transferase (GST) or a peptide epitope tag, such as FLAG or 6-His, permitting rapid, single-step, affinity-based purification of recombinant fusion protein from crude cell lysates. GST, a 26-kilodalton enzyme from Schistosoma japonicum, enables the purification of fusion proteins on immobilized glutathione under conditions that maintain protein activity and antigenicity (Amersham Pharmacia Biotech). Following purification, the GST moiety can be proteolytically cleaved from HCYT at specifically engineered sites. FLAG, an 8-amino acid peptide, enables immunoaffinity purification using commercially available monoclonal and polyclonal anti-FLAG antibodies (Eastman Kodak). 6-His, a stretch of six consecutive histidine residues, enables purification on metal-chelate resins (QIAGEN). Methods for protein expression and purification are discussed in Ausubel (1995, supra, ch 10 and 16). Purified HCYT obtained by these methods can be used directly in the following activity assay.

X. Demonstration of HCYT Activity

HCYT activity may be measured by effects of the proteins on cellular locomotion. In vitro cell motility (locomotion) assays are performed as follows. Myosin is diluted to 200 µg/ml in buffer C (25 mM imidazole, pH 7.4, 25 mM KCl, 4 mM $MgCl_2$, 1 mM EGTA, 10 mM dithiothreitol), applied to a flow cell coated with nitrocellulose, and blocked with buffer C containing 0.5 mg/ml BSA (C/BSA). A solution of phalloidin-labeled actin is perfused followed by 1 mM ATP in C/BSA to remove myosin heads that bind actin in a rigor fashion. After washing with C/BSA to remove the excess nonfluorescent actin, a solution of rhodamine-phalloidin-labeled actin and HCYT in C/BSA is introduced. Active movement is initiated at room temperature by introducing C/BSA containing 1 mM ATP and oxygen scavenger enzymes. Images (recorded using a Zeiss standard microscope (Zeiss, New York NY) equipped with a Hamamatsu SIT camera) of moving myotubes are tracked for up to 30 sec, and translocation velocities calculated using the myotube centroids to establish initial and final positions for 2 sec or 4 sec samples during the continuous movement.

Alternatively, an assay for HCYT activity measures the binding affinity of HCYT for actin as described by Hammell, R.L. and Hitchcock-DeGregori, S.E. (1997, J. Biol. Chem. 272:22409-22416). HCYT and actin are prepared from in vitro recombinant cDNA expression systems and the N-terminus of

HCYT is acetylated using methods well known in the art. Binding of N-terminal acetyl-HCYT to actin is measured by cosedimentation at 25°C in a Beckman model TL-100 centrifuge as described. The bound and free HCYT are determined by quantitative densitometry of SDS-polyacrylamide gels stained with Coomassie Blue. Apparent binding constants (K_{app}) and Hill coefficients (H) are determined by using methods well known in the art to fit the data to the equation as described by Hammell and Hitchcock-DeGregori (1997, supra). The HCYT:actin ratio determined using densitometry is normalized. Hammell and Hitchcock-DeGregori (1997, supra) have shown that saturation of binding corresponds to a HCYT:actin molar ratio of 0.14, a stoichiometry of 1 HCYT:7 actin.

XI. Functional Assays

HCYT function is assessed by expressing the sequences encoding HCYT at physiologically elevated levels in mammalian cell culture systems. cDNA is subcloned into a mammalian expression vector containing a strong promoter that drives high levels of cDNA expression. Vectors of choice include pCMV SPORT (Life Technologies) and pCR3.1 (Invitrogen, Carlsbad CA), both of which contain the cytomegalovirus promoter. 5-10 μg of recombinant vector are transiently transfected into a human cell line, preferably of endothelial or hematopoietic origin, using either liposome formulations or electroporation. 1-2 μg of an additional plasmid containing sequences encoding a marker protein are co-transfected. Expression of a marker protein provides a means to distinguish transfected cells from nontransfected cells and is a reliable predictor of cDNA expression from the recombinant vector. Marker proteins of choice include, e.g., Green Fluorescent Protein (GFP; Clontech), CD64, or a CD64-GFP fusion protein. Flow cytometry (FCM), an automated, laser optics-based technique, is used to identify transfected cells expressing GFP or CD64-GFP, and to evaluate properties, for example, their apoptotic state. FCM detects and quantifies the uptake of fluorescent molecules that diagnose events preceding or coincident with cell death. These events include changes in nuclear DNA content as measured by staining of DNA with propidium iodide; changes in cell size and granularity as measured by forward light scatter and 90 degree side light scatter; down-regulation of DNA synthesis as measured by decrease in bromodeoxyuridine uptake; alterations in expression of cell surface and intracellular proteins as measured by reactivity with specific antibodies; and alterations in plasma membrane composition as measured by the binding of fluorescein-conjugated Annexin V protein to the cell surface. Methods in flow cytometry are discussed in Ormerod, M. G. (1994) Flow Cytometry, Oxford, New York NY.

The influence of HCYT on gene expression can be assessed using highly purified populations of cells transfected with sequences encoding HCYT and either CD64 or CD64-GFP. CD64 and CD64-GFP are expressed on the surface of transfected cells and bind to conserved regions of human immunoglobulin G (IgG). Transfected cells are efficiently separated from nontransfected cells using magnetic beads coated with either human IgG or antibody against CD64 (DYNAL, Lake Success NY). mRNA can be purified from the cells using methods well known by those of skill in the art. Expression of mRNA

encoding HCYT and other genes of interest can be analyzed by northern analysis or microarray techniques.

XII. Production of HCYT Specific Antibodies

HCYT substantially purified using polyacrylamide gel electrophoresis (PAGE; see, e.g., Harrington, M.G. (1990) Methods Enzymol. 182:488-495), or other purification techniques, is used to immunize rabbits and to produce antibodies using standard protocols.

Alternatively, the HCYT amino acid sequence is analyzed using LASERGENE software (DNASTAR) to determine regions of high immunogenicity, and a corresponding oligopeptide is synthesized and used to raise antibodies by means known to those of skill in the art. Methods for selection of appropriate epitopes, such as those near the C-terminus or in hydrophilic regions are well described in the art. (See, e.g., Ausubel, 1995, supra, ch. 11.)

Typically, oligopeptides 15 residues in length are synthesized using an ABI 431A Peptide Synthesizer (Perkin-Elmer) using fmoc-chemistry and coupled to KLH (Sigma-Aldrich, St. Louis MO) by reaction with N-maleimidobenzoyl-N-hydroxysuccinimide ester (MBS) to increase immunogenicity. (See, e.g., Ausubel, 1995, supra.) Rabbits are immunized with the oligopeptide-KLH complex in complete Freund's adjuvant. Resulting antisera are tested for antipeptide activity by, for example, binding the peptide to plastic, blocking with 1% BSA, reacting with rabbit antisera, washing, and reacting with radio-iodinated goat anti-rabbit IgG.

XIII. Purification of Naturally Occurring HCYT Using Specific Antibodies

Naturally occurring or recombinant HCYT is substantially purified by immunoaffinity chromatography using antibodies specific for HCYT. An immunoaffinity column is constructed by covalently coupling anti-HCYT antibody to an activated chromatographic resin, such as CNBr-activated SEPHAROSE (Amersham Pharmacia Biotech). After the coupling, the resin is blocked and washed according to the manufacturer's instructions.

Media containing HCYT are passed over the immunoaffinity column, and the column is washed under conditions that allow the preferential absorbance of HCYT (e.g., high ionic strength buffers in the presence of detergent). The column is eluted under conditions that disrupt antibody/HCYT binding (e.g., a buffer of pH 2 to pH 3, or a high concentration of a chaotrope, such as urea or thiocyanate ion), and HCYT is collected.

XIV. Identification of Molecules Which Interact with HCYT

HCYT, or biologically active fragments thereof, are labeled with ^{125}I Bolton-Hunter reagent. (See, e.g., Bolton et al. (1973) Biochem. J. 133:529.) Candidate molecules previously arrayed in the wells of a multi-well plate are incubated with the labeled HCYT, washed, and any wells with labeled HCYT complex are assayed. Data obtained using different concentrations of HCYT are used to calculate values for the number, affinity, and association of HCYT with the candidate molecules.

Various modifications and variations of the described methods and systems of the invention will

be apparent to those skilled in the art without departing from the scope and spirit of the invention. Although the invention has been described in connection with specific preferred embodiments, it should be understood that the invention as claimed should not be unduly limited to such specific embodiments. Indeed, various modifications of the described modes for carrying out the invention which are obvious to those skilled in molecular biology or related fields are intended to be within the scope of the following claims.

Table 1

Protein SEQ ID NO:	Nucleotide SEQ ID NO:	Clone ID	Library	Fragments
1	9	1274060	TESTTUT02	1900209T6 (BLADTUT06), 1909056F6 (CONNTUT01), SAEA01408F1, SAEA02983R1
2	10	1577078	LNODNOT03	1577078H1 (LNODNOT03), 684408R6 (UTRSNOT02), 2478839T6 (SMCANOT03), 1909056F6 (CONNTUT01)
3	11	1426711	SINTBST01	1426711H1 (SINTBST01), 3749812H1 (UTRSNOT18), 3094943H1 (CERVNOT03), 2859695H1 (SININOT03), 4705751H1 (SMCRTXT01)
4	12	1676756	STOMFET01	760671R1 (BRAITUT02), 1676756H1 and 1676756T6 (STOMFET01), 2080585F6 (UTRSNOT08), 2364126F6, 2364126H1, and 2364126T6 (ADRENOT07), 2735685H1 (OVARNOT09), 2811007H1 (OVARNOT10)
5	13	1843770	COLNNOT08	262277R6 and 262277T6 (HNT2AGT01), 1509335F1 (LUNGNOT14), 1620333F6 (BRAITUT13), 1843770H1, 1843770X11, 1843770X13, 1843770X14, and 1843770X21 (COLNNOT08)
6	14	3768043	BRSTNOT24	288738H1 (EOSIHET02), 1328455F1 (PANCNOT07), 1340936F1 and 1340936T1 (COLNTUT03), 1695275F6 (COLNNOT23), 2106351R6 (BRAITUT03), 2378482F6 and 2378482T6 (ISLTNOT01), 3768043H1 (BRSTNOT24), 3988936H1 (LUNGNON03)
7	15	1655208	PROSTUT08	749902H1 (BRAITUT01), 755160H1 (BRAITUT02), 1457865H1 (COLNFET02), 1486091H1 (CORPNOT02), 1574402T6 (LNODNOT03), 1595969H1 (BRAINOT14), 1655208H1 (PROSTUT08), 1705982F6, 1705982H1, 1706039F6, and 1706039T6 (DUODNOT02), 1732192F6 (BRSTTUT08), 1850069H1 (LUNGFET03), 2363256X303D1 (LUNGFET05), 3036180H1 (BONEUNT01), 3093057H1 (BRSTNOT19), 4737159H1 (THYMNOR02)
8	16	2195418	THP1NOT01	024620R1 (ADENINB01), 1454159H1 (PENITUT01), 1797859R6 (PROSTUT05), 1943527T6 (HIPONOT01), 2195418H1 (THP1NOT01)

Table 2

SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequence	Identification	Analytical Methods
1	284	S79 T53 S215 T237 T252 T277 T74 S123 S211 S229	N199	L43-L64 L50-L71 L232-E240	tropomyosin	Blast BLOCKS PRINTS PROFILESCAN Pfam Motifs
2	158	S89 T111 T126 T151 S85 S103	N73	L106-E114	tropomyosin	Blast BLOCKS PRINTS PROFILESCAN Pfam Motifs
3	208	S53 S174 S123			tropomyosin	Blast BLOCKS PRINTS PROFILESCAN Pfam
4	156	S82 T86 T13 S36 S64		K38 to N52 E54 to L102	tropomyosin	BLOCKS PRINTS
5	876	S385 T25 T106 T168 T236 S271 T273 S329 T345 S387 S463 S479 T593 T621 S654 T665 S711 S859 T172 S324 T560 S682 T778 T808 S848 Y735 Y835	N104 N343	E199 to V247 E223 to Q260	tropomyosin, receptors, liprin (coiled-coil LAR-interacting protein)	Blast BLOCKS PRINTS Motifs

Table 2, cont'd.

SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequence	Identification	Analytical Methods
6	806	T403 S110 S111 S138 S140 S161 S176 T247 T287 T288 S299 S343 T348 T371 S417 T491 T541 S579 S615 S659 S664 T668 S29 S62 S65 S183 T197 S222 S327 T432 S633 S663 Y37 Y408	N136 N325 N666	E133 to S181 E413 to S450	tropomyosin, receptors, meningioma expressed antigen 6	Blast BLOCKS PRINTS Motifs
7	2442	S157 S697 S112 S196 T198 S272 S363 S377 S538 T544 S576 S581 S631 T662 T708 S764 S772 S781 S820 T842 T896 T910 T923 S929 T937 S1060 S1161 S1195 S1199 T1274 T1276 S1302 S1397 T1429 S1487 S1634 S1691 S1774 S1803 T1853 S2095 S2132 S2189 S2200 S2234 S2240 S2252 S2298 T2334 T2380 T2405 S38 T186 S235 S295 T396 T586 T817 S1025 S1076 T1383 T1439 T1519 S1540 T1585 T1701 S1951 S2128 S2138 T2218 S2267 T2411 Y1043	N233 N319 N626 N1352	Leucine Zipper: L309 to L330 L433 to L496 L580 to L622 L994 to L1015 L2275 to L2296 Tektin Signature: E1747 to H1763	centrosomal protein	Blast PRINTS Motifs
8	153	S42 S91 S7 T148	N122		p16-Arc	Blast

Table 3

SEQ ID NO:	Tissue Expression (Fraction of Total)	Disease Class (Fraction of Total)	Vector
9	Hematopoietic/Immune (0.216) Cardiovascular (0.189) Reproductive (0.162)	Cancer (0.378) Inflammation (0.378) Fetal (0.216)	pINCY
10	Hematopoietic/Immune (0.250) Cardiovascular (0.188) Developmental (0.125)	Inflammation (0.406) Cancer (0.375) Fetal (0.219)	pINCY
11	Reproductive (0.371) Gastrointestinal (0.236) Cardiovascular (0.157)	Cancer (0.521) Inflammation (0.164) Trauma (0.136)	pINCY
12	Nervous (0.267) Hematopoietic/Immune (0.150) Cardiovascular (0.133) Reproductive (0.133)	Cancer (0.500) Inflammation (0.366) Fetal (0.117)	pINCY
13	Nervous (0.233) Gastrointestinal (0.150) Reproductive (0.167)	Cancer (0.600) Inflammation (0.283) Fetal (0.133)	PSPORT1
14	Reproductive (0.283) Gastrointestinal (0.185) Hematopoietic/Immune (0.120)	Cancer (0.551) Inflammation (0.357) Fetal (0.212)	pINCY
15	Reproductive (0.243) Nervous (0.162) Developmental (0.135)	Cancer (0.514) Fetal (0.270) Inflammation (0.216)	pINCY
16	Nervous (0.270) Reproductive (0.254) Hematopoietic/Immune (0.159)	Cancer (0.365) Inflammation (0.365) Fetal (0.222)	pINCY

Table 4

Nucleotide SEQ ID NO:	Clone ID	Library	Library Comment
9	1274060	TESTTUT02	Library was constructed using RNA isolated from testicular tumor removed from a 31-year-old Caucasian male during unilateral orchiectomy. Pathology indicated embryonal carcinoma.
10	1577078	LNODNOT03	Library was constructed using RNA isolated from lymph node tissue obtained from a 67-year-old Caucasian male during a segmental lung resection and bronchoscopy. On microscopic exam, this tissue was found to be extensively necrotic with 10% viable tumor. Pathology for the associated tumor tissue indicated invasive grade 3-4 squamous cell carcinoma. Patient history included hemangioma. Family history included atherosclerotic coronary artery disease, benign hypertension, congestive heart failure, atherosclerotic coronary artery disease.
11	1426711	SINTBST01	Library was constructed using RNA isolated from ileum tissue obtained from an 18-year-old Caucasian female during bowel anastomosis. Pathology indicated Crohn's disease of the ileum, involving 15 cm of the small bowel. Family history included cerebrovascular disease and atherosclerotic coronary artery disease.
12	1676756	STOMFET01	Library was constructed using RNA isolated from the stomach tissue of a Caucasian female fetus, who died at 20 weeks gestation.
13	1843770	COLNNOT08	Library was constructed using RNA isolated from colon tissue removed from a 60-year-old Caucasian male during a left hemicolectomy.
14	3768043	BRSTNOT24	Library was constructed using RNA isolated from diseased breast tissue removed from a 46-year-old Caucasian female during bilateral subcutaneous mammectomy. Pathology indicated nonproliferative fibrocystic disease. Family history included breast cancer and cardiovascular disease.

Table 4, cont'd.

Nucleotide SEQ ID NO:	Clone ID	Library	Library Comment
15	1655208	PROSTUT08	Library was constructed using RNA isolated from prostate tumor tissue removed from a 60-year-old Caucasian male during radical prostatectomy and regional lymph node excision. Pathology indicated an adenocarcinoma (Gleason grade 3+4). Adenofibromatous hyperplasia was also present. The patient presented with elevated prostate specific antigen (PSA). Family history included tuberculosis, cerebrovascular disease, and arteriosclerotic coronary artery disease.
16	2195418	THP1NOT01	Library was constructed using RNA isolated from untreated THP-1 cells. THP-1 (ATCC TIB 202) is a human promonocyte line derived from the peripheral blood of a 1-year-old Caucasian male with acute monocytic leukemia (Int. J. Cancer (1980) 26:171).

Table 5

Program	Description	Reference	Parameter Threshold
ABI FACTURA	A program that removes vector sequences and masks ambiguous bases in nucleic acid sequences.	Perkin-Elmer Applied Biosystems, Foster City, CA.	
ABI/PARACEL FDF	A Fast Data Finder useful in comparing and annotating amino acid or nucleic acid sequences.	Perkin-Elmer Applied Biosystems, Foster City, CA; Paracel Inc., Pasadena, CA.	Mismatch <50%
ABI AutoAssembler	A program that assembles nucleic acid sequences.	Perkin-Elmer Applied Biosystems, Foster City, CA.	
BLAST	A Basic Local Alignment Search Tool useful in sequence similarity search for amino acid and nucleic acid sequences. BLAST includes five functions: blastp, blastn, blastx, tblastn, and tblastx.	Altschul, S.F. et al. (1990) J. Mol. Biol. 215:403-410; Altschul, S.F. et al. (1997) Nucleic Acids Res. 25: 3389-3402.	*ESTs:* Probability value= 1.0E-8 or less *Full Length sequences:* Probability value= 1.0E-10 or less
FASTA	A Pearson and Lipman algorithm that searches for similarity between a query sequence and a group of sequences of the same type. FASTA comprises as least five functions: fasta, tfasta, fastx, tfastx, and ssearch.	Pearson, W.R. and D.J. Lipman (1988) Proc. Natl. Acad Sci. 85:2444-2448; Pearson, W.R. (1990) Methods Enzymol. 183: 63-98; and Smith, T.F. and M. S. Waterman (1981) Adv. Appl. Math. 2:482-489.	*ESTs:* fasta E value=1.06E-6 *Assembled ESTs*: fasta Identity= 95% or greater and Match length=200 bases or greater; fastx E value=1.0E-8 or less *Full Length sequences:* fastx score=100 or greater
BLIMPS	A BLocks IMProved Searcher that matches a sequence against those in BLOCKS and PRINTS databases to search for gene families, sequence homology, and structural fingerprint regions.	Henikoff, S and J.G. Henikoff, Nucl. Acid Res., 19:6565-72, 1991. J.G. Henikoff and S. Henikoff (1996) Methods Enzymol. 266:88-105; and Attwood, T.K. et al. (1997) J. Chem. Inf. Comput. Sci. 37: 417-424.	Score=1000 or greater; Ratio of Score/Strength = 0.75 or larger; and Probability value= 1.0E-3 or less
PFAM	A Hidden Markov Models-based application useful for protein family search.	Krogh, A. et al. (1994) J. Mol. Biol., 235:1501-1531; Sonnhammer, E.L.L. et al. (1988) Nucleic Acids Res. 26:320-322.	Score=10-50 bits, depending on individual protein families

Table 5 cont.

Program	Description	Reference	Parameter Threshold
ProfileScan	An algorithm that searches for structural and sequence motifs in protein sequences that match sequence patterns defined in Prosite.	Gribskov, M. et al. (1988) CABIOS 4:61-66; Gribskov, et al. (1989) Methods Enzymol. 183:146-159; Bairoch, A. et al. (1997) Nucleic Acids Res. 25: 217-221.	Score= 4.0 or greater
Phred	A base-calling algorithm that examines automated sequencer traces with high sensitivity and probability.	Ewing, B. et al. (1998) Genome Res. 8:175-185; Ewing, B. and P. Green (1998) Genome Res. 8:186-194.	
Phrap	A Phils Revised Assembly Program including SWAT and CrossMatch, programs based on efficient implementation of the Smith-Waterman algorithm, useful in searching sequence homology and assembling DNA sequences.	Smith, T.F. and M. S. Waterman (1981) Adv. Appl. Math. 2:482-489; Smith, T.F. and M. S. Waterman (1981) J. Mol. Biol. 147:195-197; and Green, P., University of Washington, Seattle, WA.	Score= 120 or greater; Match length= 56 or greater
Consed	A graphical tool for viewing and editing Phrap assemblies	Gordon, D. et al. (1998) Genome Res. 8:195-202.	
SPScan	A weight matrix analysis program that scans protein sequences for the presence of secretory signal peptides.	Nielson, H. et al. (1997) Protein Engineering 10:1-6; Claverie, J.M. and S. Audic (1997) CABIOS 12: 431-439.	Score=5 or greater
Motifs	A program that searches amino acid sequences for patterns that matched those defined in Prosite.	Bairoch et al. supra; Wisconsin Package Program Manual, version 9, page M51-59, Genetics Computer Group, Madison, WI.	

What is claimed is:

1. A substantially purified polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, and SEQ ID NO:8 and fragments thereof.

2. A substantially purified variant having at least 90% amino acid sequence identity to the amino acid sequence of claim 1.

3. An isolated and purified polynucleotide encoding the polypeptide of claim 1.

4. An isolated and purified polynucleotide variant having at least 70% polynucleotide sequence identity to the polynucleotide of claim 3.

5. An isolated and purified polynucleotide which hybridizes under stringent conditions to the polynucleotide of claim 3.

6. An isolated and purified polynucleotide having a sequence which is complementary to the polynucleotide of claim 3.

7. A method for detecting a polynucleotide, the method comprising the steps of:

(a) hybridizing the polynucleotide of claim 6 to at least one nucleic acid in a sample, thereby forming a hybridization complex; and

(b) detecting the hybridization complex, wherein the presence of the hybridization complex correlates with the presence of the polynucleotide in the sample.

8. The method of claim 7 further comprising amplifying the polynucleotide prior to hybridization.

9. An isolated and purified polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:9-16 and fragments thereof.

10. An isolated and purified polynucleotide variant having at least 70% polynucleotide sequence identity to the polynucleotide of claim 9.

11. An isolated and purified polynucleotide having a sequence which is complementary to the polynucleotide of claim 9.

12. An expression vector comprising at least a fragment of the polynucleotide of claim 3.

13. A host cell comprising the expression vector of claim 12.

14. A method for producing a polypeptide, the method comprising the steps of:

a) culturing the host cell of claim 13 under conditions suitable for the expression of the polypeptide; and

b) recovering the polypeptide from the host cell culture.

15. A pharmaceutical composition comprising the polypeptide of claim 1 in conjunction with a suitable pharmaceutical carrier.

16. A purified antibody which specifically binds to the polypeptide of claim 1.

17. A purified agonist of the polypeptide of claim 1.

18. A purified antagonist of the polypeptide of claim 1.

19. A method for treating or preventing a disorder associated with decreased expression or activity of HCYT, the method comprising administering to a subject in need of such treatment an effective amount of the pharmaceutical composition of claim 15.

20. A method for treating or preventing a disorder associated with increased expression or activity of HCYT, the method comprising administering to a subject in need of such treatment an effective amount of the antagonist of claim 18.

M·H

PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7 :
C12N 15/12, C07K 14/47, 16/18, A61K 38/16, C12N 15/11, C12Q 1/68, C12N 15/62, G01N 33/566

A2

(11) International Publication Number: **WO 00/06730**

(43) International Publication Date: 10 February 2000 (10.02.00)

(21) International Application Number: PCT/US99/17167

(22) International Filing Date: 30 July 1999 (30.07.99)

(30) Priority Data:

60/155,185	31 July 1998 (31.07.98)	US
60/160,081	4 August 1998 (04.08.98)	US
60/155,228	19 August 1998 (19.08.98)	US

(63) Related by Continuation (CON) or Continuation-in-Part (CIP) to Earlier Applications

US	60/155,185 (CIP)
Filed on	31 July 1998 (31.07.98)
US	60/160,081 (CIP)
Filed on	4 August 1998 (04.08.98)
US	60/155,228 (CIP)
Filed on	19 August 1998 (19.08.98)

(71) Applicant *(for all designated States except US)*: INCYTE PHARMACEUTICALS, INC. [US/US]; 3174 Porter Drive, Palo Alto, CA 94304 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: BANDMAN, Olga [US/US]; 366 Anna Avenue, Mountain View, CA 94043 (US). TANG, Y., Tom [US/US]; 4230 Ranwick Court, San Jose, CA 95118 (US). YUE, Henry [US/US]; 826 Lois Avenue, Sunnyvale, CA 94087 (US). CORLEY, Neil, C. [US/US]; 1240 Dale Avenue #30, Mountain View, CA 94040 (US). GUEGLER, Karl, J. [US/US]; 1048 Orakland Road, Menlo Park, CA 94025 (US). AZIMZAI, Yalda [US/US]; 2045 Rock Springs Drive, Hayward, CA 94545 (US). PATTERSON, Chandra [US/US]; 490 Sherwood Way #1, Menlo Park, CA 94025 (US). LAL, Preeti [US/US]; 2382 Lass Drive, Santa Clara, CA 95054 (US). BAUGHN, Mariah, R. [US/US]; 14244 Santiago Road, San Leandro, CA 94577 (US).

(74) Agents: BILLINGS, Lucy, J. et al.; Incyte Pharmaceuticals, Inc., 3174 Porter Drive, Palo Alto, CA 94304 (US).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published
Without international search report and to be republished upon receipt of that report.

(54) Title: HUMAN CYTOSKELETAL PROTEINS

(57) Abstract

The invention provides human cytoskeletal proteins (HCYT) and polynucleotides which identify and encode HCYT. The invention also provides expression vectors, host cells, antibodies, agonists, and antagonists. The invention also provides methods for diagnosing, treating or preventing disorders associated with expression of HCYT.

```
1    MARNTLSS-RFRRVDIDEFDENKFVDEQEE  2195418
1    MSKNTVSSARFRKVDVDEYDENKFVDEED-  GI 2282042

30   AAAAAAEPGPDPSEVDGLLRQGDMLRAFHA  2195418
30   --GGDGQAGPDEGEVDSCLRQGNMTALQA   GI 2282042

60   ALRNSPVNTKNQAVKERAQGVVLKVLTNFK  2195418
58   ALKNPPINTKSQAVKDRAGSIVLKVLISFK  GI 2282042

90   SSEIEQAVQSLDRNGVDLLMKYIYKGFEKP  2195418
88   ANDIEKAVQSLDKNGVDLLMKYIYKGFESP  GI 2282042

120  TENSSAVLLQWHEKALAVGGLGSIIRVLTA  2195418
118  SDNSSAMLLQWHEKALAGGVGSIVRVLTA   GI 2282042

150  RKTV                            2195418
148  RKTV                            GI 2282042
```

FIGURE 1

Docket No.: PF-0568 USN

DECLARATION AND POWER OF ATTORNEY FOR UNITED STATES PATENT APPLICATION

As a below named inventor, I hereby declare that:

My residence, post office address and citizenship are as stated below next to my name, and

I believe that I am the original, first and sole inventor (if only one name is listed below) or an original, first and joint inventor (if more than one name is listed below) of the subject matter which is claimed and for which a United States patent is sought on the invention entitled

HUMAN CYTOSKELETAL PROTEINS

the specification of which:

/___/ is attached hereto.

/ X / was filed on January 18, 2001 as application Serial No. 09/744,314 and if this box contains an X /___/, was amended on _____.

/ X / was filed as Patent Cooperation Treaty international application No. PCT/US99/17167 on 30 July 1999, if this box contains an X /_/, was amended on under Patent Cooperation Treaty Article 19 on _____ 2001, and if this box contains an X /_/, was amended on _____.

I hereby state that I have reviewed and understand the contents of the above-identified specification, including the claims, as amended by any amendment referred to above.

I acknowledge my duty to disclose information which is material to the examination of this application in accordance with Title 37, Code of Federal Regulations, §1.56(a).

I hereby claim the benefit under Title 35, United States Code, §119 or §365(a)-(b) of any foreign application(s) for patent or inventor's certificate indicated below and of any Patent Cooperation Treaty international applications(s) designating at least one country other than the United States indicated below and have also identified below any foreign application(s) for patent or inventor's certificate and Patent Cooperation Treaty international application(s) designating at least one country other than the United States for the same subject matter and having a filing date before that of the application for said subject matter the priority of which is claimed:

Docket No.: PF-0568 USN

Country	Number	Filing Date	Priority Claimed
			/_/ Yes /_/ No
			/_/ Yes /_/ No

I hereby claim the benefit under Title 35, United States Code, §119(e) of any United States provisional application(s) listed below.

Application Serial No.	Filed	Status (Pending, Abandoned, Patented)
60/155,185	31 July 1998	Expired
60/160,081	04 August 1998	Expired
60/155,228	19 August 1998	Expired

I hereby claim the benefit under Title 35, United States Code, §120 of any United States application(s) listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in said prior application(s) in the manner required by the first paragraph of Title 35, United States Code §112, I acknowledge my duty to disclose material information as defined in Title 37 Code of Federal Regulations, §1.56(a) which occurred between the filing date(s) of the prior application(s) and the national or Patent Cooperation Treaty international filing date of this application:

Application Serial No.	Filed	Status (Pending, Abandoned, Patented)

I hereby appoint the following:

13

Lucy J. Billings	Reg. No. 36,749
Michael C. Cerrone	Reg. No. 39,132
Diana Hamlet-Cox	Reg. No. 33,302
Richard C. Ekstrom	Reg. No. 37,027
Barrie D. Greene	Reg. No. 46,740
Matthew R. Kaser	Reg. No. 44,817
Lynn E. Murry	Reg. No. 42,918
Shirley A. Recipon	Reg. No. 47,016
Susan K. Sather	Reg. No. 44,316
Michelle M. Stempien	Reg. No. 41,327
David G. Streeter	Reg. No. 43,168
Stephen Todd	Reg. No. 47,139
P. Ben Wang	Reg. No. 41,420

respectively and individually, as my patent attorneys and/or agents, with full power of substitution and revocation, to prosecute this application and to transact all business in the Patent and Trademark Office connected therewith. Please address all communications to:

LEGAL DEPARTMENT
INCYTE GENOMICS, INC.
3160 PORTER DRIVE, PALO ALTO, CA 94304

TEL: 650-855-0555 FAX: 650-849-8886 or 650-845-4166

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issuing thereon.

Sole Inventor or First Joint Inventor: 100

Full name:	OLGA BANDMAN
Signature:	Olga Bandman
Date:	16 February , 2001
Citizenship	United States
Residence:	Mountain View, California CA
P.O. Address:	366 Anna Avenue Mountain View, California 94043

Second Joint Inventor: 200

Full name:	Y. TOM TANG
Signature:	Y. Tom Tang
Date:	February 27 , 2001
Citizenship	~~People's Republic of China~~ USA
Residence:	San Jose, California CA
P.O. Address:	4230 Ranwick Court San Jose, California 95118

Y. Tom Tang 2/27/2001

Third Joint Inventor:

300

Full name:	**HENRY YUE**
Signature:	
Date:	**March 05, 2001**
Citizenship	**United States**
Residence:	**Sunnyvale, California** CA
P.O. Address:	**826 Lois Avenue** **Sunnyvale, California 94087**

Fourth Joint Inventor:

400

Full name:	**NEIL C. CORLEY**
Signature:	
Date:	**March 5, 2001**
Citizenship	**United States**
Residence:	**Castro Valley, California** CA
P.O. Address:	**20426 Crow Creek Road** **Castro Valley, California 94552**

Fifth Joint Inventor:

500

Full name:	**KARL J. GUEGLER**
Signature:	
Date:	**02/02, 2001**
Citizenship	**Switzerland**
Residence:	**Menlo Park, California** CA
P.O. Address:	**1048 Oakland Road** **Menlo Park, California 94025**

Sixth Joint Inventor:

6-00

Full name:	**YALDA AZIMZAI**
Signature:	Yalda Azimzai
Date:	February 2nd , 2001
Citizenship	**United States**
Residence:	**Castro Valley, California** CA
P.O. Address:	**5518 Boulder Canyon Drive Castro Valley, California 94552**

Seventh Joint Inventor: 7-00

Full name:	**CHANDRA PATTERSON**
Signature:	Chandra Patterson
Date:	February 7 , 2001
Citizenship	**United States**
Residence:	**Menlo Park, California** CA
P.O. Address:	**490 Sherwood Way, #1 Menlo Park, California 94025**

Eighth Joint Inventor:

8-00

Full name:	**PREETI LAL**
Signature:	Preeti Lal
Date:	FEBRUARY, 16 , 2001
Citizenship	**India**
Residence:	**Santa Clara, California** CA
P.O. Address:	**P.O. Box 5142 Santa Clara, California 95056**

Ninth Joint Inventor:

9-00

Full name:	**MARIAH R. BAUGHN**
Signature:	Mariah R Baughn
Date:	February 12 , 2001
Citizenship	**United States**
Residence:	**San Leandro, California** CA
P.O. Address:	**14244 Santiago Road San Leandro, California 94577**

SEQUENCE LISTING

<110> INCYTE PHARMACEUTICALS, INC.
BANDMAN, Olga
TANG, Y. Tom
YUE, Henry
CORLEY, Neil C.
GUEGLER, Karl J.
AZIMZAI, Yalda
PATTERSON, Chandra
LAL, Preeti
BAUGHN, Mariah R.

<120> HUMAN CYTOSKELETAL PROTEINS

<130> PF-0568 PCT

<140> To Be Assigned
<141> Herewith

<150> 09/127,665
<151> 1998-07-31

<160> 17

<170> PERL Program

<210> 1
<211> 284
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte Clone No: 1274060

<400> 1
Met Glu Ala Ile Lys Lys Lys Met Gln Met Leu Lys Leu Asp Lys
1 5 10 15
Glu Asn Ala Ile Asp Arg Ala Glu Gln Ala Glu Ala Asp Lys Lys
20 25 30
Ala Ala Glu Asp Lys Cys Lys Gln Val Glu Glu Glu Leu Thr His
35 40 45
Leu Gln Lys Lys Leu Lys Gly Thr Glu Asp Glu Leu Asp Lys Tyr
50 55 60
Ser Glu Asp Leu Lys Asp Ala Gln Glu Lys Leu Glu Leu Thr Glu
65 70 75
Lys Lys Ala Ser Asp Ala Glu Gly Asp Val Ala Ala Leu Asn Arg
80 85 90
Arg Ile Gln Leu Val Glu Glu Glu Leu Asp Arg Ala Gln Glu Arg
95 100 105
Leu Ala Thr Ala Leu Gln Lys Leu Glu Glu Ala Glu Lys Ala Ala
110 115 120
Asp Glu Ser Glu Arg Gly Met Lys Val Ile Glu Asn Arg Ala Met
125 130 135
Lys Asp Glu Glu Lys Met Glu Ile Gln Glu Met Gln Leu Lys Glu
140 145 150
Ala Lys His Ile Ala Glu Glu Ala Asp Arg Lys Tyr Glu Glu Val
155 160 165
Ala Arg Lys Leu Val Ile Leu Glu Gly Glu Leu Glu Arg Ala Glu
170 175 180
Glu Arg Ala Glu Val Ser Glu Leu Lys Cys Gly Asp Leu Glu Glu
185 190 195

```
Glu Leu Lys Asn Val Thr Asn Asn Leu Lys Ser Leu Glu Ala Ala
                200                 205                 210
Ser Glu Lys Tyr Ser Glu Lys Glu Asp Lys Tyr Glu Glu Glu Ile
                215                 220                 225
Lys Leu Leu Ser Asp Lys Leu Lys Glu Ala Glu Thr Arg Ala Glu
                230                 235                 240
Phe Ala Glu Arg Thr Val Ala Lys Leu Glu Lys Thr Ile Asp Asp
                245                 250                 255
Leu Glu Glu Lys Leu Ala Gln Ala Lys Glu Glu Asn Val Gly Leu
                260                 265                 270
His Gln Thr Leu Asp Gln Thr Leu Asn Glu Leu Asn Cys Ile
                275                 280
```

<210> 2
<211> 158
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte Clone No: 1577078

<400> 2

```
Met Lys Val Ile Glu Asn Arg Ala Met Lys Asp Glu Glu Lys Met
  1               5                  10                  15
Glu Ile Gln Glu Met Gln Leu Lys Glu Ala Lys His Ile Ala Glu
            20                  25                  30
Glu Ala Asp Arg Lys Tyr Glu Glu Val Ala Arg Lys Leu Val Ile
            35                  40                  45
Leu Glu Gly Glu Leu Glu Arg Ala Glu Glu Arg Ala Glu Val Ser
            50                  55                  60
Glu Leu Lys Cys Gly Asp Leu Glu Glu Glu Leu Lys Asn Val Thr
            65                  70                  75
Asn Asn Leu Lys Ser Leu Glu Ala Ala Ser Glu Lys Tyr Ser Glu
            80                  85                  90
Lys Glu Asp Lys Tyr Glu Glu Glu Ile Lys Leu Leu Ser Asp Lys
            95                 100                 105
Leu Lys Glu Ala Glu Thr Arg Ala Glu Phe Ala Glu Arg Thr Val
           110                 115                 120
Ala Lys Leu Glu Lys Thr Ile Asp Asp Leu Glu Glu Lys Leu Ala
           125                 130                 135
Gln Ala Lys Glu Glu Asn Val Gly Leu His Gln Thr Leu Asp Gln
           140                 145                 150
Thr Leu Asn Glu Leu Asn Cys Ile
           155
```

<210> 3
<211> 208
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte Clone No: 1426711

<400> 3

```
Met Asp Ala Ile Lys Lys Met Gln Met Leu Lys Leu Asp Lys
 1               5                  10                  15
Glu Asn Ala Leu Asp Arg Ala Glu Gln Ala Glu Ala Asp Lys Lys
            20                  25                  30
Ala Ala Glu Asp Arg Ser Lys Gln Leu Glu Glu Asp Ile Ala Ala
        35                  40                  45
Lys Glu Lys Leu Leu Arg Val Ser Glu Asp Glu Arg Asp Arg Val
    50                  55                  60
Leu Glu Glu Leu His Lys Ala Glu Asp Ser Leu Leu Ala Ala Glu
65                  70                  75
Glu Ala Ala Ala Lys Ala Glu Ala Asp Val Ala Ser Leu Asn Arg
                80                  85                  90
Arg Ile Gln Leu Val Glu Glu Glu Leu Asp Arg Ala Gln Glu Arg
            95                  100                 105
Leu Ala Thr Ala Leu Gln Lys Leu Glu Glu Ala Glu Lys Ala Ala
        110                 115                 120
Asp Glu Ser Glu Arg Gly Met Lys Val Ile Glu Ser Arg Ala Gln
    125                 130                 135
Lys Asp Glu Glu Lys Met Glu Ile Gln Glu Ile Gln Leu Lys Glu
    140                 145                 150
Ala Lys His Ile Ala Glu Asp Ala Asp Arg Lys Tyr Glu Glu Val
155                 160                 165
Ala Arg Lys Leu Val Ile Ile Glu Ser Asp Leu Glu Arg Ala Glu
                170                 175                 180
Glu Arg Ala Gly Glu Gly Leu Asp Lys Asp Arg Arg Ala Ala Thr
            185                 190                 195
His His Pro Ser Pro His Pro His Pro Leu Leu Glu Phe
        200                 205
```

<210> 4
<211> 156
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte Clone No: 1676756

<400> 4

```
Met Ile Lys Arg Val Leu Leu Glu Arg Leu Glu Asn Thr Arg Lys
 1               5                  10                  15
Leu Arg Glu Leu Thr Glu Gly Arg Thr Leu Asp Trp Pro Gln Asn
            20                  25                  30
Arg Ile Thr Glu Val Ser Ala Lys Arg Gln Ile Val Thr Glu Tyr
        35                  40                  45
Arg Glu Lys Gly Lys Arg Asn Tyr Glu Glu Lys Lys Arg Asp Leu
    50                  55                  60
Glu Gly Arg Ser Arg Arg Tyr Asn Leu Cys Ile Ile Gly Ile Pro
65                  70                  75
Glu Thr Glu Asp Arg Ala Ser Gly Ala Glu Thr Ile Lys Asp Leu
                80                  85                  90
Leu Glu Lys Asn Phe Pro Glu Leu Lys Asn Glu Leu Asp Leu Gln
            95                  100                 105
Met Glu Lys Ala His Arg Ile Pro Leu Lys Phe Asn Glu Lys Lys
        110                 115                 120
Ala Ala Ser Arg His Ile Arg Val Thr Phe Leu Asn Phe Lys Asp
    125                 130                 135
Glu Thr Phe Tyr Lys His Pro Val Arg Glu Ser Arg Leu Leu Thr
    140                 145                 150
Lys Gly Gln Lys Ser Gly
155
```

<210> 5
<211> 876
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte Clone No: 1843770

<400> 5
Met Ala Ser Asp Ala Ser His Ala Leu Glu Ala Ala Leu Glu Gln
1 5 10 15
Met Asp Gly Ile Ile Ala Gly Thr Lys Thr Gly Ala Asp Leu Ser
20 25 30
Asp Gly Thr Cys Glu Pro Gly Leu Ala Ser Pro Ala Ser Tyr Met
35 40 45
Asn Pro Phe Pro Val Leu His Leu Ile Glu Asp Leu Arg Leu Ala
50 55 60
Leu Glu Met Leu Glu Leu Pro Gln Glu Arg Ala Ala Leu Leu Ser
65 70 75
Gln Ile Pro Gly Pro Thr Ala Ala Tyr Ile Lys Glu Trp Phe Glu
80 85 90
Glu Ser Leu Ser Gln Val Asn His His Ser Ala Ala Ser Asn Glu
95 100 105
Thr Tyr Gln Glu Arg Leu Ala Arg Leu Glu Gly Asp Lys Glu Ser
110 115 120
Leu Ile Leu Gln Val Ser Val Leu Thr Asp Gln Val Glu Ala Gln
125 130 135
Gly Glu Lys Ile Arg Asp Leu Glu Val Cys Leu Glu Gly His Gln
140 145 150
Val Lys Leu Asn Ala Ala Glu Glu Met Leu Gln Gln Glu Leu Leu
155 160 165
Ser Arg Thr Ser Leu Glu Thr Gln Lys Leu Asp Leu Met Thr Glu
170 175 180
Val Ser Glu Leu Lys Leu Lys Leu Val Gly Met Glu Lys Glu Gln
185 190 195
Arg Glu Gln Glu Glu Lys Gln Arg Lys Ala Glu Glu Leu Leu Gln
200 205 210
Glu Leu Arg His Leu Lys Ile Lys Val Glu Glu Leu Glu Asn Glu
215 220 225
Arg Asn Gln Tyr Glu Trp Lys Leu Lys Ala Thr Lys Ala Glu Val
230 235 240
Ala Gln Leu Gln Glu Gln Val Ala Leu Lys Asp Ala Glu Ile Glu
245 250 255
Arg Leu His Ser Gln Leu Ser Arg Thr Ala Ala Leu His Ser Glu
260 265 270
Ser His Thr Glu Arg Asp Gln Glu Ile Gln Arg Leu Lys Met Gly
275 280 285
Met Glu Thr Leu Leu Leu Ala Asn Glu Asp Lys Asp Arg Arg Ile
290 295 300
Glu Glu Leu Thr Gly Leu Leu Asn Gln Tyr Arg Lys Val Lys Glu
305 310 315
Ile Val Met Val Thr Gln Gly Pro Ser Glu Arg Thr Leu Ser Ile
320 325 330
Asn Glu Glu Glu Pro Glu Gly Gly Phe Ser Lys Trp Asn Ala Thr
335 340 345
Asn Lys Asp Pro Glu Glu Leu Phe Lys Gln Glu Met Pro Pro Arg
350 355 360
Cys Ser Ser Pro Thr Val Gly Pro Pro Pro Leu Pro Gln Lys Ser
365 370 375
Leu Glu Thr Arg Ala Gln Lys Lys Leu Ser Cys Ser Leu Glu Asp
380 385 390
Leu Arg Ser Glu Ser Val Asp Lys Cys Met Asp Gly Asn Gln Pro
395 400 405

Phe Pro Val Leu Glu Pro Lys Asp Ser Pro Phe Leu Ala Glu His
410 415 420
Lys Tyr Pro Thr Leu Pro Gly Lys Leu Ser Gly Ala Thr Pro Asn
425 430 435
Gly Glu Ala Ala Lys Ser Pro Pro Thr Ile Cys Gln Pro Asp Ala
440 445 450
Thr Gly Ser Ser Leu Leu Arg Leu Arg Asp Thr Glu Ser Gly Trp
455 460 465
Asp Asp Thr Ala Val Val Asn Asp Leu Ser Ser Thr Ser Ser Gly
470 475 480
Thr Glu Ser Gly Pro Gln Ser Pro Leu Thr Pro Asp Gly Lys Arg
485 490 495
Asn Pro Lys Gly Ile Lys Lys Phe Trp Gly Lys Ile Arg Arg Thr
500 505 510
Gln Ser Gly Asn Phe Tyr Thr Asp Thr Leu Gly Met Ala Glu Phe
515 520 525
Arg Arg Gly Gly Leu Arg Ala Thr Ala Gly Pro Arg Leu Ser Arg
530 535 540
Thr Arg Asp Ser Lys Gly Gln Lys Ser Asp Ala Asn Ala Pro Phe
545 550 555
Ala Gln Trp Ser Thr Glu Arg Val Cys Ala Trp Leu Glu Asp Phe
560 565 570
Gly Leu Ala Gln Tyr Val Ile Phe Ala Arg Gln Trp Val Ser Ser
575 580 585
Gly His Thr Leu Leu Thr Ala Thr Pro Gln Asp Met Glu Lys Glu
590 595 600
Leu Gly Ile Lys His Pro Leu His Arg Lys Lys Leu Val Leu Ala
605 610 615
Val Lys Ala Ile Asn Thr Lys Gln Glu Glu Lys Ser Ala Leu Leu
620 625 630
Asp His Ile Trp Val Thr Arg Trp Leu Asp Asp Ile Gly Leu Pro
635 640 645
Gln Tyr Lys Asp Gln Phe His Glu Ser Arg Val Asp Gly Arg Met
650 655 660
Leu Gln Tyr Leu Thr Val Asn Asp Leu Leu Phe Leu Lys Val Thr
665 670 675
Ser Gln Leu His His Leu Ser Ile Lys Cys Ala Ile His Val Leu
680 685 690
His Val Asn Lys Phe Asn Pro His Cys Leu His Arg Arg Pro Ala
695 700 705
Asp Glu Ser Asn Leu Ser Pro Ser Glu Val Val Gln Trp Ser Asn
710 715 720
His Arg Val Met Glu Trp Leu Arg Ser Val Asp Leu Ala Glu Tyr
725 730 735
Ala Pro Asn Leu Arg Gly Ser Gly Val His Gly Gly Leu Ile Ile
740 745 750
Leu Glu Pro Arg Phe Thr Gly Asp Thr Leu Ala Met Leu Leu Asn
755 760 765
Ile Pro Pro Gln Lys Thr Leu Leu Arg Arg His Leu Thr Thr Lys
770 775 780
Phe Asn Ala Leu Ile Gly Pro Glu Ala Glu Gln Glu Lys Arg Glu
785 790 795
Lys Met Ala Ser Pro Ala Tyr Thr Pro Leu Thr Thr Thr Ala Lys
800 805 810
Val Arg Pro Arg Lys Leu Gly Phe Ser His Phe Gly Asn Ile Arg
815 820 825
Lys Lys Lys Phe Asp Glu Ser Thr Asp Tyr Ile Cys Pro Met Glu
830 835 840
Pro Ser Asp Gly Val Ser Asp Ser His Arg Val Tyr Ser Gly Tyr
845 850 855
Arg Gly Leu Ser Pro Leu Asp Ala Pro Glu Leu Asp Gly Leu Asp
860 865 870
Gln Val Gly Gln Ile Ser
875

<210> 6
<211> 806
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte Clone No: 3768043

<400> 6
Met Glu Glu Pro Gly Ala Thr Pro Gln Pro Tyr Leu Gly Leu Val
1 5 10 15
Leu Glu Glu Leu Arg Arg Val Val Ala Ala Leu Pro Glu Ser Met
20 25 30
Arg Pro Asp Glu Asn Pro Tyr Gly Phe Pro Ser Glu Leu Val Val
35 40 45
Cys Ala Ala Val Ile Gly Phe Phe Val Val Leu Leu Phe Leu Trp
50 55 60
Arg Ser Phe Arg Ser Val Arg Ser Arg Leu Tyr Val Gly Arg Glu
65 70 75
Gln Lys Leu Gly Ala Thr Leu Ser Gly Leu Ile Glu Glu Lys Cys
80 85 90
Lys Leu Leu Glu Lys Phe Ser Leu Ile Gln Lys Glu Tyr Glu Gly
95 100 105
Tyr Glu Val Glu Ser Ser Leu Glu Asp Ala Ser Phe Glu Lys Glu
110 115 120
Ala Ala Glu Glu Ala Arg Ser Leu Glu Ala Thr Cys Glu Lys Leu
125 130 135
Asn Arg Ser Asn Ser Glu Leu Glu Asp Glu Ile Leu Cys Leu Glu
140 145 150
Lys Asp Leu Lys Gln Glu Lys Ser Lys His Ser Gln Gln Asp Glu
155 160 165
Leu Met Ala Asp Ile Ser Lys Ser Ile Gln Ser Leu Glu Asp Glu
170 175 180
Ser Lys Ser Leu Lys Ser Gln Ile Ala Glu Ala Lys Ile Ile Cys
185 190 195
Lys Thr Phe Lys Met Ser Glu Glu Arg Arg Ala Ile Ala Ile Lys
200 205 210
Asp Ala Leu Asn Glu Asn Ser Gln Leu Gln Thr Ser His Lys Gln
215 220 225
Leu Phe Gln Gln Glu Ala Glu Val Trp Lys Gly Glu Val Ser Glu
230 235 240
Leu Asn Lys Gln Lys Ile Thr Phe Glu Asp Ser Lys Val His Ala
245 250 255
Glu Gln Val Leu Asn Asp Lys Glu Asn His Ile Lys Thr Leu Thr
260 265 270
Gly His Leu Pro Met Met Lys Asp Gln Ala Ala Val Leu Glu Glu
275 280 285
Asp Thr Thr Asp Asp Asp Asn Leu Glu Leu Glu Val Asn Ser Glu
290 295 300
Ser Glu Asn Gly Ala Tyr Leu Asp Asn Pro Pro Lys Gly Ala Leu
305 310 315
Lys Lys Leu Ile His Ala Ala Lys Leu Asn Ala Ser Leu Lys Thr
320 325 330
Leu Glu Gly Glu Arg Asn Gln Ile Tyr Ile Gln Leu Ser Glu Val
335 340 345
Asp Lys Thr Lys Glu Glu Leu Thr Glu His Ile Lys Asn Leu Gln
350 355 360
Thr Gln Gln Ala Ser Leu Gln Ser Glu Asn Thr His Phe Glu Asn
365 370 375
Glu Asn Gln Lys Leu Gln Gln Lys Leu Lys Val Met Thr Glu Leu
380 385 390

```
Tyr Gln Glu Asn Glu Met Lys Leu His Arg Lys Leu Thr Val Glu
                395                 400                 405
Glu Asn Tyr Arg Leu Glu Lys Glu Glu Lys Leu Ser Lys Val Asp
            410                 415                 420
Glu Lys Ile Ser His Ala Thr Glu Glu Leu Glu Thr Tyr Arg Lys
        425                 430                 435
Arg Ala Lys Asp Leu Glu Glu Glu Leu Glu Arg Thr Ile His Ser
    440                 445                 450
Tyr Gln Gly Gln Ile Ile Ser His Glu Lys Lys Ala His Asp Asn
455                 460                 465
Trp Leu Ala Ala Arg Asn Ala Glu Arg Asn Leu Asn Asp Leu Arg
            470                 475                 480
Lys Glu Asn Ala His Asn Arg Gln Lys Leu Thr Glu Thr Glu Leu
        485                 490                 495
Lys Phe Glu Leu Leu Glu Lys Asp Pro Tyr Ala Leu Asp Val Pro
    500                 505                 510
Asn Thr Ala Phe Gly Arg Glu His Ser Pro Tyr Gly Pro Ser Pro
515                 520                 525
Leu Gly Trp Pro Ser Ser Glu Thr Arg Ala Phe Leu Ser Pro Pro
            530                 535                 540
Thr Leu Leu Glu Gly Pro Leu Arg Leu Ser Pro Leu Leu Pro Gly
        545                 550                 555
Gly Gly Gly Arg Gly Ser Arg Gly Pro Gly Asn Pro Leu Asp His
    560                 565                 570
Gln Ile Thr Asn Glu Arg Gly Glu Ser Ser Cys Asp Arg Leu Thr
575                 580                 585
Asp Pro His Arg Ala Pro Ser Asp Thr Gly Ser Leu Ser Pro Pro
            590                 595                 600
Trp Asp Gln Asp Arg Arg Met Met Phe Pro Pro Pro Gly Gln Ser
        605                 610                 615
Tyr Pro Asp Ser Ala Leu Pro Pro Gln Arg Gln Asp Arg Phe Cys
    620                 625                 630
Ser Asn Ser Gly Arg Leu Ser Gly Pro Ala Glu Leu Arg Ser Phe
635                 640                 645
Asn Met Pro Ser Leu Asp Lys Met Asp Gly Ser Met Pro Ser Glu
            650                 655                 660
Met Glu Ser Ser Arg Asn Asp Thr Lys Asp Asp Leu Gly Asn Leu
        665                 670                 675
Asn Val Pro Asp Ser Ser Leu Pro Ala Glu Asn Glu Ala Thr Gly
    680                 685                 690
Pro Gly Phe Val Pro Pro Pro Leu Ala Pro Ile Arg Gly Pro Leu
695                 700                 705
Phe Pro Val Asp Ala Arg Gly Pro Phe Leu Arg Arg Gly Pro Pro
            710                 715                 720
Phe Pro Pro Pro Pro Pro Gly Ala Met Phe Gly Ala Ser Arg Asp
        725                 730                 735
Tyr Phe Pro Pro Arg Asp Phe Pro Gly Pro Pro Pro Ala Pro Phe
    740                 745                 750
Ala Met Arg Asn Val Tyr Pro Pro Arg Gly Phe Pro Pro Tyr Leu
755                 760                 765
Pro Pro Arg Pro Gly Phe Phe Pro Pro Pro Pro His Ser Glu Gly
            770                 775                 780
Arg Ser Glu Phe Pro Ser Gly Leu Ile Pro Pro Ser Asn Glu Pro
        785                 790                 795
Ala Thr Glu His Pro Glu Pro Gln Gln Glu Thr
    800                 805
```

<210> 7
<211> 2442
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte Clone No: 1655208

<400> 7
Met Glu Thr Arg Ser Pro Gly Leu Asn Asn Met Lys Pro Gln Ser
1 5 10 15
Leu Gln Leu Val Leu Glu Glu Gln Val Leu Ala Leu Gln Gln Gln
20 25 30
Met Ala Glu Asn Gln Ala Ala Ser Trp Arg Lys Leu Lys Asn Ser
35 40 45
Gln Glu Ala Gln Gln Arg Gln Ala Thr Leu Val Arg Lys Leu Gln
50 55 60
Ala Lys Val Leu Gln Tyr Arg Ser Trp Cys Gln Glu Leu Glu Lys
65 70 75
Arg Leu Glu Ala Thr Gly Gly Pro Ile Pro Gln Arg Trp Glu Asn
80 85 90
Val Glu Glu Pro Asn Leu Asp Glu Leu Leu Val Arg Leu Glu Glu
95 100 105
Glu Gln Gln Arg Cys Glu Ser Leu Ala Glu Val Asn Thr Gln Ile
110 115 120
Arg Leu His Met Glu Lys Ala Asp Val Val Asn Lys Ala Leu Arg
125 130 135
Ala Asp Val Glu Lys Leu Thr Val Asp Trp Ser Arg Ala Arg Asp
140 145 150
Glu Leu Met Arg Lys Glu Ser Gln Trp Gln Met Glu Gln Glu Phe
155 160 165
Phe Lys Gly Tyr Leu Lys Gly Glu His Gly Arg Leu Leu Ser Leu
170 175 180
Trp Arg Glu Val Val Thr Phe Arg Arg His Phe Leu Glu Met Lys
185 190 195
Ser Ala Thr Asp Arg Asp Leu Met Glu Leu Lys Ala Glu His Val
200 205 210
Arg Leu Ser Gly Ser Leu Leu Thr Cys Cys Leu Arg Leu Thr Val
215 220 225
Gly Ala Gln Ser Arg Glu Pro Asn Gly Ser Gly Arg Met Asp Gly
230 235 240
Arg Glu Pro Ala Gln Leu Leu Leu Leu Leu Ala Lys Thr Gln Glu
245 250 255
Leu Glu Lys Glu Ala His Glu Arg Ser Gln Glu Leu Ile Gln Leu
260 265 270
Lys Ser Gln Gly Asp Leu Glu Lys Ala Glu Leu Gln Asp Arg Val
275 280 285
Thr Glu Leu Ser Ala Leu Leu Thr Gln Ser Gln Lys Gln Asn Glu
290 295 300
Asp Tyr Glu Lys Met Ile Lys Ala Leu Arg Glu Thr Val Glu Ile
305 310 315
Leu Glu Thr Asn His Thr Glu Leu Met Glu His Glu Ala Ser Leu
320 325 330
Ser Arg Asn Ala Gln Glu Glu Lys Leu Ser Leu Gln Gln Val Ile
335 340 345
Lys Asp Ile Thr Gln Val Met Val Glu Glu Gly Asp Asn Ile Ala
350 355 360
Gln Gly Ser Gly Leu Glu Asn Ser Leu Glu Leu Glu Ser Ser Ile
365 370 375
Phe Ser Gln Phe Asp Tyr Gln Asp Ala Asp Lys Ala Leu Thr Leu
380 385 390
Val Arg Ser Val Leu Thr Arg Arg Arg Gln Ala Val Gln Asp Leu
395 400 405
Arg Gln Gln Leu Ala Gly Cys Gln Glu Ala Val Asn Leu Leu Gln
410 415 420
Gln Gln His Asp Gln Trp Glu Glu Glu Gly Lys Ala Leu Arg Gln
425 430 435
Arg Leu Gln Lys Leu Thr Gly Glu Arg Asp Thr Leu Ala Gly Gln
440 445 450

Thr Val Asp Leu Gln Gly Glu Val Asp Ser Leu Ser Lys Glu Arg
455 460 465
Glu Leu Leu Gln Lys Ala Arg Glu Glu Leu Arg Gln Gln Leu Glu
470 475 480
Val Leu Glu Gln Glu Ala Trp Arg Leu Arg Arg Val Asn Val Glu
485 490 495
Leu Gln Leu Gln Gly Asp Ser Ala Gln Gly Gln Lys Glu Glu Gln
500 505 510
Gln Glu Glu Leu His Leu Ala Val Arg Glu Arg Glu Arg Leu Gln
515 520 525
Glu Met Leu Met Gly Leu Glu Ala Lys Gln Ser Glu Ser Leu Ser
530 535 540
Glu Leu Ile Thr Leu Arg Glu Ala Leu Glu Ser Ile His Leu Glu
545 550 555
Gly Glu Leu Leu Arg Gln Glu Gln Thr Glu Val Thr Ala Ala Leu
560 565 570
Ala Arg Ala Glu Gln Ser Ile Ala Glu Leu Ser Ser Ser Glu Asn
575 580 585
Thr Leu Lys Thr Glu Val Ala Asp Leu Arg Ala Ala Ala Val Lys
590 595 600
Leu Ser Ala Leu Asn Glu Ala Leu Ala Leu Asp Lys Val Gly Leu
605 610 615
Asn Gln Gln Leu Leu Gln Leu Glu Glu Glu Asn Gln Ser Val Cys
620 625 630
Ser Arg Met Glu Ala Ala Glu Gln Ala Arg Asn Ala Leu Gln Val
635 640 645
Asp Leu Ala Glu Ala Glu Lys Arg Arg Glu Ala Leu Trp Glu Lys
650 655 660
Asn Thr His Leu Glu Ala Gln Leu Gln Lys Ala Glu Glu Ala Gly
665 670 675
Ala Glu Leu Gln Ala Asp Leu Arg Asp Ile Gln Glu Glu Lys Glu
680 685 690
Glu Ile Gln Lys Lys Leu Ser Glu Ser Arg His Gln Gln Glu Ala
695 700 705
Ala Thr Thr Gln Leu Glu Gln Leu His Gln Glu Ala Lys Arg Gln
710 715 720
Glu Glu Val Leu Ala Arg Ala Val Gln Glu Lys Glu Ala Leu Val
725 730 735
Arg Glu Lys Ala Ala Leu Glu Val Arg Leu Gln Ala Val Glu Arg
740 745 750
Asp Arg Gln Asp Leu Ala Ala Gln Leu Gln Gly Leu Ser Ser Ala
755 760 765
Lys Glu Leu Leu Glu Ser Ser Leu Phe Glu Ala Gln Gln Gln Asn
770 775 780
Ser Val Ile Asp Glu Pro Gln Gly Gln Leu Glu Val Gln Ile Gln
785 790 795
Thr Val Thr Gln Ala Lys Glu Val Ile Gln Gly Glu Val Arg Cys
800 805 810
Leu Lys Leu Glu Leu Asp Thr Glu Arg Ser Gln Ala Glu Gln Glu
815 820 825
Arg Asp Ala Ala Ala Arg Gln Leu Ala Gln Ala Glu Gln Glu Gly
830 835 840
Lys Thr Ala Leu Glu Gln Gln Lys Ala Ala His Glu Lys Glu Val
845 850 855
Asn Gln Leu Arg Glu Lys Trp Glu Lys Glu Arg Ser Trp His Gln
860 865 870
Gln Glu Leu Ala Lys Ala Leu Glu Ser Leu Glu Arg Glu Lys Met
875 880 885
Glu Leu Glu Met Arg Leu Lys Glu Gln Gln Thr Glu Met Glu Ala
890 895 900
Ile Gln Ala Gln Arg Glu Glu Glu Arg Thr Gln Ala Glu Ser Ala
905 910 915
Leu Cys Gln Met Gln Leu Glu Thr Glu Lys Glu Arg Val Ser Leu
920 925 930
Leu Glu Thr Leu Leu Gln Thr Gln Lys Glu Leu Ala Asp Ala Ser

```
                935                 940                 945
Gln Gln Leu Glu Arg Leu Arg Gln Asp Met Lys Val Gln Lys Leu
                950                 955                 960
Lys Glu Gln Glu Thr Thr Gly Ile Leu Gln Thr Gln Leu Gln Glu
                965                 970                 975
Ala Gln Arg Glu Leu Lys Glu Ala Ala Arg Gln His Arg Asp Asp
                980                 985                 990
Leu Ala Ala Leu Gln Glu Glu Ser Ser Ser Leu Leu Gln Asp Lys
                995                1000                1005
Met Asp Leu Gln Lys Gln Val Glu Asp Leu Lys Ser Gln Leu Val
               1010                1015                1020
Ala Gln Asp Asp Ser Gln Arg Leu Val Glu Gln Glu Val Gln Glu
               1025                1030                1035
Lys Leu Arg Glu Thr Gln Glu Tyr Asn Arg Ile Gln Lys Glu Leu
               1040                1045                1050
Glu Arg Glu Lys Ala Ser Leu Thr Leu Ser Leu Met Glu Lys Glu
               1055                1060                1065
Gln Arg Leu Leu Val Leu Gln Glu Ala Asp Ser Ile Arg Gln Gln
               1070                1075                1080
Glu Leu Ser Ala Leu Arg Gln Asp Met Gln Glu Ala Gln Gly Glu
               1085                1090                1095
Gln Lys Glu Leu Ser Ala Gln Met Glu Leu Leu Arg Gln Glu Val
               1100                1105                1110
Lys Glu Lys Glu Ala Asp Phe Leu Ala Gln Glu Ala Gln Leu Leu
               1115                1120                1125
Glu Glu Leu Glu Ala Ser His Ile Thr Glu Gln Gln Leu Arg Ala
               1130                1135                1140
Ser Leu Trp Ala Gln Glu Ala Lys Ala Ala Gln Leu His Leu Arg
               1145                1150                1155
Leu Arg Ser Thr Glu Ser Gln Leu Glu Ala Leu Ala Ala Glu Gln
               1160                1165                1170
Gln Pro Gly Asn Gln Ala Gln Ala Gln Ala Gln Leu Ala Ser Leu
               1175                1180                1185
Tyr Ser Ala Leu Gln Gln Ala Leu Gly Ser Val Cys Glu Ser Arg
               1190                1195                1200
Pro Glu Leu Ser Gly Gly Gly Asp Ser Ala Pro Ser Val Trp Gly
               1205                1210                1215
Leu Glu Pro Asp Gln Asn Gly Ala Arg Ser Leu Phe Lys Arg Gly
               1220                1225                1230
Pro Leu Leu Thr Ala Leu Ser Ala Glu Ala Val Ala Ser Ala Leu
               1235                1240                1245
Leu Lys Leu His Gln Asp Leu Trp Lys Thr Gln Gln Thr Arg Asp
               1250                1255                1260
Val Leu Arg Asp Gln Val Gln Lys Leu Glu Glu Arg Leu Thr Asp
               1265                1270                1275
Thr Glu Ala Glu Lys Ser Gln Val His Thr Glu Leu Gln Asp Leu
               1280                1285                1290
Gln Arg Gln Leu Ser Gln Asn Gln Glu Glu Lys Ser Lys Trp Glu
               1295                1300                1305
Gly Lys Gln Asn Ser Leu Glu Ser Glu Leu Met Glu Leu His Glu
               1310                1315                1320
Thr Met Ala Ser Leu Gln Ser Arg Leu Arg Arg Ala Glu Leu Gln
               1325                1330                1335
Arg Met Glu Ala Gln Gly Glu Arg Glu Leu Leu Gln Ala Ala Lys
               1340                1345                1350
Glu Asn Leu Thr Ala Gln Val Glu His Leu Gln Ala Ala Val Val
               1355                1360                1365
Glu Ala Arg Ala Gln Ala Ser Ala Ala Gly Ile Leu Glu Glu Asp
               1370                1375                1380
Leu Arg Thr Ala Arg Ser Ala Leu Lys Leu Lys Asn Glu Glu Val
               1385                1390                1395
Glu Ser Glu Arg Glu Arg Ala Gln Ala Leu Gln Glu Gln Gly Glu
               1400                1405                1410
Leu Lys Val Ala Gln Gly Lys Ala Leu Gln Glu Asn Leu Ala Leu
               1415                1420                1425
```

Leu Thr Gln Thr Leu Ala Glu Arg Glu Glu Glu Val Glu Thr Leu
1430 1435 1440
Arg Gly Gln Ile Gln Glu Leu Glu Lys Gln Arg Glu Met Gln Lys
1445 1450 1455
Ala Ala Leu Glu Leu Leu Ser Leu Asp Leu Lys Lys Arg Asn Gln
1460 1465 1470
Glu Val Asp Leu Gln Gln Glu Gln Ile Gln Glu Leu Glu Lys Cys
1475 1480 1485
Arg Ser Val Leu Glu His Leu Pro Met Ala Val Gln Glu Arg Glu
1490 1495 1500
Gln Lys Leu Thr Val Gln Arg Glu Gln Ile Arg Glu Leu Glu Lys
1505 1510 1515
Asp Arg Glu Thr Gln Arg Asn Val Leu Glu His Gln Leu Leu Glu
1520 1525 1530
Leu Glu Lys Lys Asp Gln Met Ile Glu Ser Gln Arg Gly Gln Val
1535 1540 1545
Gln Asp Leu Lys Lys Gln Leu Val Thr Leu Glu Cys Leu Ala Leu
1550 1555 1560
Glu Leu Glu Glu Asn His His Lys Met Glu Cys Gln Gln Lys Leu
1565 1570 1575
Ile Lys Glu Leu Glu Gly Gln Arg Glu Thr Gln Arg Val Ala Leu
1580 1585 1590
Thr His Leu Thr Leu Asp Leu Glu Glu Arg Ser Gln Glu Leu Gln
1595 1600 1605
Ala Gln Ser Ser Gln Ile His Asp Leu Glu Ser His Ser Thr Val
1610 1615 1620
Leu Ala Arg Glu Leu Gln Glu Arg Asp Gln Glu Val Lys Ser Gln
1625 1630 1635
Arg Glu Gln Ile Glu Glu Leu Gln Arg Gln Lys Glu His Leu Thr
1640 1645 1650
Gln Asp Leu Glu Arg Arg Asp Gln Glu Leu Met Leu Gln Lys Glu
1655 1660 1665
Arg Ile Gln Val Leu Glu Asp Gln Arg Thr Arg Gln Thr Lys Ile
1670 1675 1680
Leu Glu Glu Asp Leu Glu Gln Ile Lys Leu Ser Leu Arg Glu Arg
1685 1690 1695
Gly Arg Glu Leu Thr Thr Gln Arg Gln Leu Met Gln Glu Arg Ala
1700 1705 1710
Glu Glu Gly Lys Gly Pro Ser Lys Ala Gln Arg Gly Ser Leu Glu
1715 1720 1725
His Met Lys Leu Ile Leu Arg Asp Lys Glu Lys Glu Val Glu Cys
1730 1735 1740
Gln Gln Glu His Ile His Glu Leu Gln Glu Leu Lys Asp Gln Leu
1745 1750 1755
Glu Gln Gln Leu Gln Gly Leu His Arg Lys Val Gly Glu Thr Ser
1760 1765 1770
Leu Leu Leu Ser Gln Arg Glu Gln Glu Ile Val Val Leu Gln Gln
1775 1780 1785
Gln Leu Gln Glu Ala Arg Glu Gln Gly Glu Leu Lys Glu Gln Ser
1790 1795 1800
Leu Gln Ser Gln Leu Asp Glu Ala Gln Arg Ala Leu Ala Gln Arg
1805 1810 1815
Asp Gln Glu Leu Glu Ala Leu Gln Gln Glu Gln Gln Gln Ala Gln
1820 1825 1830
Gly Gln Glu Glu Arg Val Lys Glu Lys Ala Asp Ala Leu Gln Gly
1835 1840 1845
Ala Leu Glu Gln Ala His Met Thr Leu Lys Glu Arg His Gly Glu
1850 1855 1860
Leu Gln Asp His Lys Glu Gln Ala Arg Arg Leu Glu Glu Glu Leu
1865 1870 1875
Ala Val Glu Gly Arg Arg Val Gln Ala Leu Glu Glu Val Leu Gly
1880 1885 1890
Asp Leu Arg Ala Glu Ser Arg Glu Gln Glu Lys Ala Leu Leu Ala
1895 1900 1905
Leu Gln Gln Gln Cys Ala Glu Gln Ala Gln Glu His Glu Val Glu

1910 1915 1920
Thr Arg Ala Leu Gln Asp Ser Trp Leu Gln Ala Gln Ala Val Leu
1925 1930 1935
Lys Glu Arg Asp Gln Glu Leu Glu Ala Leu Arg Ala Glu Ser Gln
1940 1945 1950
Ser Ser Arg His Gln Glu Glu Ala Ala Arg Ala Arg Ala Glu Ala
1955 1960 1965
Leu Gln Glu Ala Leu Gly Lys Ala His Ala Ala Leu Gln Gly Lys
1970 1975 1980
Glu Gln His Leu Leu Glu Gln Ala Glu Leu Ser Arg Ser Leu Glu
1985 1990 1995
Ala Ser Thr Ala Thr Leu Gln Ala Ser Leu Asp Ala Cys Gln Ala
2000 2005 2010
His Ser Arg Gln Leu Glu Glu Ala Leu Arg Ile Gln Glu Gly Glu
2015 2020 2025
Ile Gln Asp Gln Asp Leu Arg Tyr Gln Glu Asp Val Gln Gln Leu
2030 2035 2040
Gln Gln Ala Leu Ala Gln Arg Asp Glu Glu Leu Arg His Gln Gln
2045 2050 2055
Glu Arg Glu Gln Leu Leu Glu Lys Ser Leu Ala Gln Arg Val Gln
2060 2065 2070
Glu Asn Met Ile Gln Glu Lys Gln Asn Leu Gly Gln Glu Arg Glu
2075 2080 2085
Glu Glu Glu Ile Arg Gly Leu His Gln Ser Val Arg Glu Leu Gln
2090 2095 2100
Leu Thr Leu Ala Gln Lys Glu Gln Glu Ile Leu Glu Leu Arg Glu
2105 2110 2115
Thr Gln Gln Arg Asn Asn Leu Glu Ala Leu Pro His Ser His Lys
2120 2125 2130
Thr Ser Pro Met Glu Glu Gln Ser Leu Lys Leu Asp Ser Leu Glu
2135 2140 2145
Pro Arg Leu Gln Arg Glu Leu Glu Arg Leu Gln Ala Ala Leu Arg
2150 2155 2160
Gln Thr Glu Ala Arg Glu Ile Glu Trp Arg Glu Lys Ala Gln Asp
2165 2170 2175
Leu Ala Leu Ser Leu Ala Gln Thr Lys Ala Ser Val Ser Ser Leu
2180 2185 2190
Gln Glu Val Ala Met Phe Leu Gln Ala Ser Val Leu Glu Arg Asp
2195 2200 2205
Ser Glu Gln Gln Arg Leu Gln Asp Glu Leu Glu Leu Thr Arg Arg
2210 2215 2220
Ala Leu Glu Lys Glu Arg Leu His Ser Pro Gly Ala Thr Ser Thr
2225 2230 2235
Ala Glu Leu Gly Ser Arg Gly Glu Gln Gly Val Gln Leu Gly Glu
2240 2245 2250
Val Ser Gly Val Glu Ala Glu Pro Ser Pro Asp Gly Met Glu Lys
2255 2260 2265
Gln Ser Trp Arg Gln Arg Leu Glu His Leu Gln Gln Ala Val Ala
2270 2275 2280
Arg Leu Glu Ile Asp Arg Ser Arg Leu Gln Arg His Asn Val Gln
2285 2290 2295
Leu Arg Ser Thr Leu Glu Gln Val Glu Arg Glu Arg Arg Lys Leu
2300 2305 2310
Lys Arg Glu Ala Met Arg Ala Ala Gln Ala Gly Ser Leu Glu Ile
2315 2320 2325
Ser Lys Ala Thr Ala Ser Ser Pro Thr Gln Gln Asp Gly Arg Gly
2330 2335 2340
Gln Lys Asn Ser Asp Ala Lys Cys Val Ala Glu Leu Gln Lys Glu
2345 2350 2355
Val Val Leu Leu Gln Ala Gln Leu Thr Leu Glu Arg Lys Gln Lys
2360 2365 2370
Gln Asp Tyr Ile Thr Arg Ser Ala Gln Thr Ser Arg Glu Leu Ala
2375 2380 2385
Gly Leu His His Ser Leu Ser His Ser Leu Leu Ala Val Ala Gln
2390 2395 2400

```
Ala Pro Glu Ala Thr Val Leu Glu Ala Glu Thr Arg Arg Leu Asp
                2405                2410                2415
Glu Ser Leu Thr Gln Ser Leu Thr Ser Pro Gly Pro Val Leu Leu
                2420                2425                2430
His Pro Ser Pro Ser Thr Thr Gln Ala Ala Ser Arg
                2435                2440
```

<210> 8
<211> 153
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte Clone No: 2195418

<400> 8
```
Met Ala Arg Asn Thr Leu Ser Ser Arg Phe Arg Arg Val Asp Ile
  1               5                  10                  15
Asp Glu Phe Asp Glu Asn Lys Phe Val Asp Glu Gln Glu Glu Ala
             20                  25                  30
Ala Ala Ala Ala Ala Glu Pro Gly Pro Asp Pro Ser Glu Val Asp
         35                  40                  45
Gly Leu Leu Arg Gln Gly Asp Met Leu Arg Ala Phe His Ala Ala
     50                  55                  60
Leu Arg Asn Ser Pro Val Asn Thr Lys Asn Gln Ala Val Lys Glu
 65                  70                  75
Arg Ala Gln Gly Val Val Leu Lys Val Leu Thr Asn Phe Lys Ser
             80                  85                  90
Ser Glu Ile Glu Gln Ala Val Gln Ser Leu Asp Arg Asn Gly Val
         95                 100                 105
Asp Leu Leu Met Lys Tyr Ile Tyr Lys Gly Phe Glu Lys Pro Thr
        110                 115                 120
Glu Asn Ser Ser Ala Val Leu Leu Gln Trp His Glu Lys Ala Leu
    125                 130                 135
Ala Val Gly Gly Leu Gly Ser Ile Ile Arg Val Leu Thr Ala Arg
140                 145                 150
Lys Thr Val
```

<210> 9
<211> 1185
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte Clone No: 1274060

<400> 9
```
cacagccaca gcccctgact gccgcagccc ccacagagcc cgccgcgcac cccacgtccc  60
ccacgccagc gcccagccat ggaggccatc aagaagaaaa tgcagatgct gaagttggac 120
aaggagaatg ccatcgaccg cgcggagcag gcggaggcgg ataagaaagc cgctgaggac 180
aagtgcaagc aggtggagga ggagctgacg cacctccaga agaaactaaa agggacagag 240
gacgagctgg ataaatattc cgaggacctg aaggacgcgc aggagaagct ggagctcacg 300
gagaagaagg cctccgacgc tgaaggtgat gtggccgccc tcaaccgacg catccagctc 360
gttgaggagg agttggacag ggctcaggaa cgactggcca cggccctgca gaagctggag 420
gaggcagaaa aagctgcaga tgagagtgag agaggaatga aggtgataga aaaccgggcc 480
atgaaggatg aggagaagat ggagattcag gagatgcagc tcaaagaggc caagcacatt 540
gcggaagagg ctgaccgcaa atacgaggag gtagctcgta agctggtcat cctggagggt 600
```

```
gagctggaga gggcagagga gcgtgcggag gtgtctgaac taaaatgtgg tgacctggaa 660
gaagaactca agaatgttac taacaatctg aaatctctgg aggctgcatc tgaaaagtat 720
tctgaaaagg aggacaaata tgaagaagaa attaaacttc tgtctgacaa actgaaagag 780
gctgagaccc gtgctgaatt tgcagagaga acggttgcaa aactggaaaa gacaattgat 840
gacctggaag agaaacttgc ccaggccaaa gaagagaacg tgggcttaca tcagacactg 900
gatcagacac taaacgaact taactgtata taagcaaaac agaagagtct tgttccaaca 960
gaaactctgg agctccgtgg gtctttctct tctcttgtaa gaagttcctt ttgttattgc 1020
catcttcgct ttgctggaaa tgtcaagcaa attatgaata catgaccaaa tattttgtat 1080
cggagaagct ttgagcacca tgttaaatct cattccttcc ctttttttt caaaaaaaa 1140
aagagatagg attgagtgga agggtagggg aggaaaaaaa aaaaa 1185
```

<210> 10
<211> 1050
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte Clone No: 1577078

<400> 10

```
gtttttccca gacattttcc gggaagatca tactgaaatt ttggtatgag tgtaaattcc  60
ctatggcctg gactcctggg tgggctttga cagggaagat caggttaaca gagggcagga 120
catgggggag gctcccactg gtggctggcc tgatacttct taacatggct gcacctccga 180
cctccccagg ctgaaggtga tgtggccgcc ctcaaccgac gcatccagct cgttgaggag 240
gagttggaca gggctcagga acgactggcc acggccctgc agaagctgga ggaggcagaa 300
aaagctgcag atgagagtga gagaggaatg aaggtgatag aaaaccgggc catgaaggat 360
gaggagaaga tggagattca ggagatgcag ctcaaagagg ccaagcacat tgcggaagag 420
gctgaccgca aatacgagga ggtagctcgt aagctggtca tcctggaggg tgagctggag 480
agggcagagg agcgtgcgga ggtgtctgaa ctaaaatgtg gtgacctgga agaagaactc 540
aagaatgtta ctaacaatct gaaatctctg gaggctgcat ctgaaaagta ttctgaaaag 600
gaggacaaat atgaagaaga aattaaactt ctgtctgaca aactgaaaga ggctgagacc 660
cgtgctgaat ttgcagagag aacggttgca aaactggaaa agacaattga tgacctggaa 720
gagaaacttg cccaggccaa agaagagaac gtgggcttac atcagacact ggatcagaca 780
ctaaacgaac ttaactgtat ataagcaaaa cagaagagtc ttgttccaac agaaactctg 840
gagctccgtg ggtctttctc ttctcttgta agaagttcct tttgttattg ccatctttcgc 900
tttgctggaa atgtcaagca aattatgaat acatgaccaa atattttgta tcggagaagc 960
tttgagcacc atgttaaatc tcattccttc ccttttttt tcaaaaaaaa aaagagatag 1020
gattgagtgg aagggtaggg gaggaaaaaa                                  1050
```

<210> 11
<211> 729
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte Clone No: 1426711

<400> 11

```
cgctcctccg cccgaccgcg cgctcgcccc gccgctcctg ctgcagcccc agggcccctc  60
gccgccgcca ccatggacgc catcaagaag aagatgcaga tgctgaagct cgacaaggag 120
aacgccttgg atcgagctga gcaggcggag gccgacaaga aggcggcgga agacaggagc 180
aagcagctcg aggaggacat cgcggccaag gagaagttgc tgcgggtgtc ggaggacgag 240
cgggaccggg tgctggagga gctgcacaag gcggaggaca gcctcctggc cgccgaagag 300
gccgccgcca aggctgaagc cgacgtagct tctctgaaca gacgcatcca gctggttgag 360
gaagagttgg atcgtgccca ggagcgtctg gcaacagctt tgcagaagct ggaggaagct 420
gagaaggcag cagatgagag tgagagaggc atgaaagtca ttgagagtcg agcccaaaaa 480
gatgaagaaa aaatggaaat tcaggagatc caactgaaag aggcaaagca cattgctgaa 540
gatgccgacc gcaaatatga agaggtggcc cgtaagctgg tcatcattga gagcgacctg 600
gaacgtgcag aggagcgggc tggagaaggc ctggataaag acagaagggc ggcaacacac 660
```

caccccagcc cccaccccca ccctctcctt gagtttctgt gaattaaaat atttgcaaat 720
ccaaaaaaa 729

<210> 12
<211> 2068
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte Clone No: 1676756

<400> 12
gcagccagaa ccgagtcagc cataaagcta cgcgctaggc tcttggccct gacgtgaggc 60
gcgcagagat ggcgtaacgg gaatagtttt caacgtctat ttcattccct gcttcagagg 120
acctctttaa tctttgattt tggtccctgt ttctaagaaa agcaactgaa aaggtcgtaa 180
taccgcccct gagaaaaaag gagcagcgct aaataatcga gaaaatgcct cctcttgaaa 240
cggatataga gatggaaaca agatataaga aggattgaga atcatataat acaggagctt 300
aaacacctat gcgcgatgat aaagagggta ctattagagc gcttggaaaa taccaggaag 360
ttgagagagt taacagaagg gcgcacgctg gattggccac aaaatcgaat tactgaagta 420
agtgcaaaac gacaaattgt cacagaatac agagaaaagg ggaaaagaaa ttacgaggag 480
aaaaagagag atctagaggg ccggtccagg agatacaatc tatgcataat aggaatacct 540
gaaactgagg acagagcaag tggagctgaa acaataaagg atctacttga aaaaaatttt 600
ccagaattga agaacgaact agatctacaa atggaaaagg ctcataggat acctttaaag 660
tttaatgaaa agaaagcagc atctagacat atccgggtga cgtttttgaa tttcaaagac 720
gaaacatttt acaagcatcc agtcagagaa agcaggttac ttacaaaggg gcaaaagtca 780
ggctgacctc agatttttct cctgcaattc taaatgccag aagacagtgg aacaatatct 840
agagtgttaa gggaaaataa ttttgagcca agaattatat actctgccaa gttatcattt 900
ctttacaaag gaaactggaa gacattctta gatatacagg ggttaggaaa gtatatcaac 960
caagaacttt ccctgaaaat tttgctgaag gatttactgc agctaacaga gaacctgaat 1020
taaaataaga atagggaggc aattgtatga aagaactgat ggtatgcatt aaaactagtt 1080
aaggagcatt aagtttaaat tgttatgtat agggcattaa aactaaatac aaaaatagaa 1140
acataagctg ggtgtaaaat ttcagtttac gtttaaactg gaagggcata ccacatgttc 1200
aatatattgg acaaatactg ttttattttt agcatgtagt attgtagata caaatacgca 1260
tgttcttaag ctaaacctaa gtaaaagggc tttgtcactg tagatttgac aaataacatt 1320
ctaaatgtca aatctgacaa ttggaaaaca gtacagtttt atttccctta tttttaagta 1380
tagcaggaaa aaaataacaa gtttgtccat gttactactg atactgtaaa gaaacttcca 1440
aatcactggt ttgggagggg gtggatacat cctcattttc tgtacatgat gcatggggaa 1500
aaacagtagt tactgcatca tatagttcca agtaaatcat atagttccaa gtaaatttga 1560
catggaaata agtgttaact acaaaaatag aatatataca ttccaatgtt tggaaatgaa 1620
gtggtaaaaa gaaaaatagc aaaaagacac taaaattagc aaaaagtaga ttgacagtac 1680
agatgaccta taaactggtt aaatagccct attacaattg aaattgggtc aaaaatgaat 1740
ctctaacctg agatacattt gtataaaaca aatgatcaaa gagttaaag ataatggaaa 1800
ctataaatca tagaaatagc aacatttttc ttaaaatcta aacataagac attttatgta 1860
aaacatacaa tataggatgt aacatttaac aaactatatg gtaaataata tgacattaat 1920
atataaaata tatgttccag atcttgatgt gggtgatagt tttacttatg tacatcgata 1980
tacatctatt taaaaacatg tacatttaaa ttgttgcatt ttatgtattt aaaatttatt 2040
gagtaatcct tctgtatgta ttattaat 2068

<210> 13
<211> 3191
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte Clone No: 1843770

<400> 13
aggcaggcaa aaggagtcat ggcttctgat gctagtcatg cgctggaagc tgccctggag 60
caaatggacg ggatcattgc aggcactaaa acaggtgcag atcttagtga tggtacttgt 120
gagcctggac tggcttcccc ggcctcctac atgaaccct tcccggtgct ccatctcatc 180

gaggacttga ggctggcctt ggagatgctg gagcttcctc aggagagagc agccctcctg 240
agccagatcc ctggcccaac agctgcctac ataaaggaat ggtttgaaga gagcttgtcc 300
caggtaaacc accacagtgc tgctagtaat gaaacctacc aggaacgctt ggcacgtcta 360
gaaggggata aggagtccct catattgcag gtgagtgtcc tcacagacca agtagaagcc 420
cagggagaaa agattcgaga cctggaagtg tgtctggaag gacaccaggt gaaactcaat 480
gctgctgaag agatgcttca acaggagctg ctaagccgca catctcttga gacccagaag 540
ctcgatctga tgactgaagt gtctgagctg aagctcaagc tggttggcat ggagaaggag 600
cagagagagc aggaggagaa gcagagaaaa gcagaggagt tactgcaaga gctcaggcac 660
ctcaaaatca aagtggaaga gttggaaaat gaaaggaatc agtatgaatg gaagctaaag 720
gccactaagg ctgaagtcgc ccagctgcaa gaacaggtgg ccctgaaaga tgcagaaatt 780
gagcgtctgc acagccagct ctcccggaca gcagctctcc acagtgagag tcacacagag 840
agagaccaag aaattcaacg tctgaaaatg gggatggaaa ctttgctgct tgccaatgaa 900
gataaggacc gtcggataga ggagcttacg gggctgttaa accagtaccg gaaggtaaag 960
gagattgtga tggtcactca agggccttcg gagagaactc tctcaatcaa tgaagaagaa 1020
ccggagggag gtttcagcaa gtggaacgct acaaataagg accctgaaga attatttaaa 1080
caagagatgc ctccaagatg tagctctcct acagtggggc cacctccatt gccacagaaa 1140
tcactggaaa ccagggctca gaaaaagctc tcttgtagtc tagaagactt gagaagtgaa 1200
tctgtggata agtgtatgga tgggaaccag cccttcccgg tgttagaacc caaggacagc 1260
cctttcttgg cggagcacaa atatcccact ttacctggga agctttcagg agccacgccc 1320
aatggagagg ctgccaaatc tcctcccacc atctgccagc ctgacgccac ggggagcagc 1380
ctgctgaggc tgagagacac agaaagtggc tgggatgaca ctgctgtggt caatgacctc 1440
tcatccacat catcgggcac tgaatcaggt cctcagtctc ctctgacacc agatggtaaa 1500
cggaatccca aaggcattaa gaagttctgg ggaaaaatcc gaagaactca gtcaggaaat 1560
ttctacactg acacgctggg gatggcagag tttcgacgag gtgggctccg ggcaaccgca 1620
gggccaagac tctctaggac cagggactcc aagggacaga aaagtgacgc caatgccccc 1680
tttgcccagt ggagcacaga gcgtgtgtgt gcatggctgc aggactttgg cctggctcag 1740
tatgtgatct ttgccaggca gtgggtatct tctggccaca ccttattgac agccaccccc 1800
caggacatgg aaaaggagct aggaattaag cacccactcc acaggaagaa gcttgtttta 1860
gcagtgaaag ccatcaacac caaacaggag gagaagtctg cactgctaga ccacatttgg 1920
gtgacaaggt ggcttgatga tattggctta ccccagtaca aagaccagtt tcatgaatct 1980
agagttgacg gacgaatgct gcaatacctа actgtgaacg atttactctt cttaaaagtc 2040
accagccaac tacatcatct cagcatcaaa tgtgccattc acgtgctgca tgtcaacaag 2100
ttcaaccccc actgcctgca ccggcggcca gctgatgaga gtaacctttc tccttcagaa 2160
gttgtacagt ggtccaacca cagggtgatg gagtggttac gatctgtgga cctggcagag 2220
tatgcaccca atcttcgagg gagtggagtc catggaggcc tcattatcct ggagccacgc 2280
ttcactgggg acaccctggc tatgcttctc aacatccccc cacaaaagac gctcctcagg 2340
cgccacctga ccaccaagtt caatgccttg attggtccgg aggctgaaca ggaaaagcga 2400
gagaaaatgg cctcaccagc ttacacacca ctgaccacca cagccaaagt ccggccaagg 2460
aaactaggat tttcacactt cggaaacata agaaaaaaga agttcgatga atcgacggac 2520
tacatttgcc caatggagcc cagtgacggt gtcagtgata gtcacaggct ctacagtggc 2580
taccggggcc tcagccccct tgatgccccт gaactggatg ggctggacca ggtgggacag 2640
attagctgat gcccttgtca cctgccctct gtgcaccctg agagctcaca gtaacactgt 2700
gtgtgtcacc atataactgc acctcacccc cgcacgtgtg catgactcgc agagaatatt 2760
ccagcaattg tgtacccctg ggccagtctc tttgaaccct gagggtggcc aggatctgga 2820
gctgcatctc taaggggcca ggctttgggg accattgcca aaggtggact caggaggaaa 2880
gacacttaaa gacactttta catgtctagt aattcttgat gttcatcttc agcaccagtg 2940
gaaacacatg aacttcgatg caggtccaga gaccatggac actcccacga ggctcagctc 3000
tcaggcaccc cctacacttc agttgaggga aaagctcaag tgccttaggc ccgtggacca 3060
cagtcttggc tgagatcaaa gggatgagca acagggactt ctgccacagt gacaatggaa 3120
ttgtgttgtg ccttacttca gaggtggtct cttctttctt gtaataaaag caatatttat 3180
gcggaaagca t 3191

<210> 14
<211> 3164
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte Clone No: 3768043

<400> 14
aggtttaatc catgaagaag acagcaattt taaaagtgta ttcaccaaaa aataaagctt 60
caaaatatgt gatgtgaaaa ctgccagaac taaggcgggc cgggctcaga ccagcgctgc 120

ctcaggatgt aaagtgtaac aagagggcca ggggaggtgg tgggggacaa catgggcctg 180
tgaggcctgt gggtgcccgc gttccccagc tcccccgcca gcccgctcca cagtggtccg 240
ctccggttgg ttgtcacgtg cgcattcggg ttccagaccc aaggctgcgt gttctccacc 300
gcttgttgtg gccagtgtta ctgcggtgac cgccagagca gcctcgacgc tatggaggag 360
cctggtgcta cccctcagcc ctacctgggg ctggtcctgg aggagctacg cagagttgtg 420
gcagcactac ctgagagtat gagaccagat gagaatcctt atggttttcc atcggaactg 480
gtggtatgtg cagctgttat tggatttttt gttgttctcc ttttttgtg gagaagtttt 540
agatcggtta ggagtcggct ttacgtggga agagagcaaa aacttggtgc aacgctttct 600
ggactaattg aagaaaaatg taaactactt gaaaaattta gccttattca aaaagagtat 660
gaaggctatg aagtagagtc atctttagag gatgccagct ttgagaagga ggcagcagaa 720
gaagcacgaa gtttggaggc aacctgtgaa aagctgaaca ggtccaattc tgaacttgag 780
gatgaaatcc tctgtctaga aaaagactta aaacaagaga aatctaaaca ttctcaacaa 840
gatgaattga tggcggatat ttcaaaaagt atacagtctc tagaagatga gtcaaaatcc 900
ctcaaatcac aaatagctga agccaaaatc atctgcaaga catttaaaat gagtgaagaa 960
cgacgggcta tagcaataaa agatgctttg aatgaaaatt ctcaacttca gacaagccat 1020
aaacagcttt ttcagcaaga agctgaagta tggaaaggag aagtgagtga acttaataaa 1080
cagaaaataa catttgaaga ctccaaagta cacgcagaac aagttctgaa tgataaagaa 1140
aatcacatca agaccctgac tggacacttg ccaatgatga aagatcaggc tgctgtgctt 1200
gaagaagaca caacggatga tgataacctg gaattagaag tgaacagtga atcggaaaat 1260
ggtgcttact tagataatcc tccaaaagga gctttgaaga aactgattca tgctgctaag 1320
ttaaatgctt ctttaaaaac cttagaagga gaaagaaacc aaatttatat tcagttgtct 1380
gaagttgata aaacaaagga agagcttaca gagcatatta aaaatcttca gactcaacaa 1440
gcatctttgc agtcagaaaa cacacatttt gaaaatgaga atcagaagct tcaacagaaa 1500
cttaaagtaa tgactgaatt atatcaagaa aatgaaatag aactccacag gaaattaaca 1560
gtagaggaaa attatcggtt agagaaagaa gagaaacttt ctaaagtaga tgaaaagatc 1620
agccatgcca ctgaagagct ggagacctat agaaagcgag ccaaagatct tgaagaagaa 1680
ttggagagaa ctattcattc ttatcaaggg cagattattt cccatgagaa aaaagcacat 1740
gataattggt tggcagctcg gaatgctgaa agaaacctca atgatttaag gaaagaaaat 1800
gctcacaaca gacaaaaatt aactgaaaca gagcttaaat ttgaactttt agaaaaagat 1860
ccttatgcac tcgatgttcc aaatacagca tttggcagag agcattcccc atatggtccc 1920
tcaccattgg gttggccttc atctgaaaca agagcttttc tctctcctcc aacttgttg 1980
gagggtccac tcagactctc acctttgctt ccagggggag gaggaagagg ctcacgaggc 2040
ccagggaatc ctctggacca tcagattacc aatgaaagag gagaatcaag ctgtgatagg 2100
ttaaccgatc ctcatagggc tccctctgac actgggtctc tgtcacctcc atgggaccag 2160
gaccgtagga tgatgtttcc tccgccagga caatcatatc ctgattcagc ccttcctcca 2220
caaaggcaag acagattttg ttctaattct ggtagactgt ctggaccagc agaactcaga 2280
agttttaata tgccttcttt ggataaaatg gatgggtcaa tgccttcaga aatggaatcc 2340
agtagaaata atacccaaaga tgatcttggt aatttaaatg tgcctgattc atctctccct 2400
gctgaaaatg aagccactgg ccctggcttt gttcctccac ctcttgctcc aatcagaggt 2460
ccattgtttc cagtggatgc aagaggccca ttcttgagaa gaggacctcc tttccccca 2520
cctcctccag gagccatgtt tggagcttct cgagattatt ttccaccaag ggatttccca 2580
ggtccaccac ctgctccatt tgcaatgaga aatgtctatc caccgagggg ttttcctcct 2640
taccttcccc caagacctgg atttttcccc ccacccccac attctgaagg tagaagtgag 2700
ttcccctcag gtttgattcc accttcaaat gagcctgcta ctgaacatcc agaaccacag 2760
caagaaacct gacaatattt ttgctctctt caaaagtaat tttgactgat ctcattttca 2820
gtttaagtaa ctgctgttac ttaagtgatt acactttgc tcaaattgaa gcttaatgga 2880
attataattc tcaggatagt attttgtaaa taaagatgat ttaaatatga atcttatgag 2940
taaattattt caatttttatt ttagacggta taactatttc aatttgatta atccactatt 3000
atataaacaa tagtgggagt tttatatatg taatctttca ggtggggagg ctttaaattc 3060
tgaagtctgt gtctttatgc caagaactgt atttactgtg gttgtggaca aatgtgaaag 3120
taactttatg cttaaataaa ttatagttga tttaaaaaaa aaaa 3164

<210> 15
<211> 7962
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte Clone No: 1655208

<400> 15
aaaagtagga aggtagagtt gttggcagaa atcctgggat aagagaatag tttcctggaa 60
gatctgtgcc tccaaccagc agagagggat tgagcttcat tgaactcaac agagccaaca 120

```
tttcatagca ccatgttcaa gaggaggttg aagtggcatg gcaatggtta gagaccctgc 180
tgggcgtgaa caccctctgg ctacctaggg acctgtgggc ctaccacctg gtgccctcat 240
ggagacaaga agccctgggt tgaacaacat gaagccccag tcactgcagc tggtactgga 300
agagcaggtg ctggcactac agcagcagat ggcagagaat caggcagcct cctggcggaa 360
gctgaagaac tcccaggagg cccagcagag acaagcaacc cttgtgagga agctgcaggc 420
caaggtgctg cagtaccgaa gctggtgcca agagctggag aagccgctag aagccaatgg 480
aggaccaatc ccccagaggt gggaaaatgt ggaggagcca aacctggatg agctgctggt 540
ccgattggag gaggagcaac agaggtgtga gagtctagca gaggtgaaca cccagatgcg 600
actgcacatg gaaaaagctg acgtggtgaa taaagccctt aggggcagatg tggaaaaact 660
gacagtggac tggagccggg cccgggatga gctaatgagg aaggagagcc agtggcagat 720
ggagcaggag ttcttcaagg gctacctgaa aggggagcac ggtcgcctcc tcagtctatg 780
gcgggaggtt gtgacattcc gacgccactt cctggaaatg aagtcagcta ctgccagaga 840
tctgatggag ctaaaagctg agcatgtgag gctttcaggg tctctgttga cctgttgtct 900
gcgcttgact gtgggagcac agtctcggga acccaacgga tctggaagaa tggatgggcg 960
ggagccggcc cagctgctgc tgctactagc caagacccag gagctggaga aggaagccca 1020
tgaaaggagc caggagttaa tacagctgaa gagtcaaggg gatctggaga aggctgaact 1080
tcaggaccgg gtgaccgagc tctctgctct gttgacccag tctcagaagc aaaatgaaga 1140
ttatgaaaag atgataaagg ctctgagaga gacagtggag atcctggaga caaatcacac 1200
agaattaatg gaacatgaag catctcttag taggaatgcg caagaggaga agttgtcttt 1260
acagcaggtg atcaaggata taacccaggt catggtggaa gaaggggaca atatagccca 1320
aggctctggt cttgagaact ctttggaatt ggagtctagt atcttctccc agtttgatta 1380
ccaagatgca gacaaggctc ttactctggt gcgttcagtg ctgactcgga gacgccaggc 1440
tgtgcaggac ctaaggcagc agcttgcagg ctgtcaagag gctgtgaact tgtgcaacca 1500
gcagcatgat cagtgggagg aagagggcaa agccttgaga cagcggctgc agaagctcac 1560
tggggagcgg gacactctgg cagggcagac tgtggacctc cagggagagg tggcactctc 1620
cagcaaggag cgagagctgc tgcagaaggc cagggaagag ctgcggcagc agctggaggt 1680
gctagagcag gaggcatggc gcctgcgaag ggtaaatgtg gagcttcagc tgcagggggа 1740
ctctgcccag ggccagaagg aggaacagca ggaggagctg cacctggctg tccgggagag 1800
ggagcgtctt caggagatgc tgatgggcct ggaagccaaa cagtcagaat cactcagtga 1860
actgatcact cttcgggaag ccctggagtc aattcacctg gaagtggagt tactgaggca 1920
agagcaaacg gaagtgaccg cagcgctggc tagggcagag cagtcaattg cagagctgtc 1980
gagttctgaa aacaccctga agacagaagt agctgatctt cgggctgcag ctgtcaagct 2040
cagtgcctta aatgaggctt tggcgttaga taaagttggg ctgaaccagc agcttctcca 2100
gttagaggag gagaaccagt ctgtgtgcag cagaatggag gcagcagagc aggcgagaaa 2160
tgctttgcag gtcgacctgg cggaggcaga gaagaggagg gaagccctgt gggaaaagaa 2220
cactcacctg gaggctcagc tgcagaaagc tgaggaggct gcagcctgagc tgcaggcaga 2280
tctcaggcac atccaagaag agaaggaaga aattcaaaag aaactaagtg agtcacgtca 2340
ccagcaggag gcagccacga ctcagctgga gcagctacat caggaggcaa agcgacagga 2400
agaagtgctt gccagggcag tccaggagaa ggaggcccta gtacgagaga aagcggctct 2460
agaggtgcgg ctgcaggccg tggagcgtga ccggcaggac ctcgctgcac aactacaggg 2520
gctcagctca gccaaggagc tactggagag cagtctgttt gaagcccaac aacaaaattc 2580
tgtgatagac gagccgcagg ggcagctgga ggtccagatt caaactgtca ctcaagccaa 2640
ggaagtaatc caaggggaag tgaggtgcct gaagctggaa ctggacactg aacggagtca 2700
ggcagagcag gagcgggatg ctgcagccag acagctggcc caggctgagc aagaagggaa 2760
gactgccttg gagcagcaga aggcagccca tgagaaagag gtgaaccagc tccggggaga 2820
atgggagaag gagcgctcct ggcaccagca ggagctggca aaggctctgg agagcttaga 2880
aagggaaaaa atggagctgg aaatgaggct aaaggagcag cagacagaaa tggaggccat 2940
ccaggcccag agggaagaag aacggaccca ggcagagagt gccctatgcc agatgcagct 3000
ggaaacagag aaggagagag tatccctcct ggagacactg ctgcagacgc agaaggagct 3060
agcagatgcc agccaacaac tggaacgact gaggcaggac atgaaagtcc agaaattaaa 3120
ggagcaggag accactggga tactacagac ccagctccag gaggctcaac gggagctgaa 3180
ggaggcagcc cggcagcaca gagatgacct tgctgccctc caagaagaga gcagctccct 3240
gctgcaggat aagatggacc tgcagaagca ggtggaggac ttgaagtctc agctggtggc 3300
ccaggatgac tcccagaggc tggtggagca ggaggttcag gagaagctga gagagaccca 3360
ggagtataac cgaattcaga aggagctgga gagagagaaa gccagcctga ctctgtcact 3420
gatggaaaag gaacagagac tccttgtttt acaagaagct gactctattc gacaacaaga 3480
gctgagtgcc ctgcgccagg acatgcagga ggcccaggga gaacagaaag agctcagtgc 3540
tcagatggaa ttactaaggc aagaggtgaa ggaaaaggag gctgactttc tggcccagga 3600
agcacagctg ctggagggagc tggaggcgtc tcatatcacg gagcagcagc tgcgagcctc 3660
cttgtgggcc caggaagcca aggcagccca actacacctg cgactgcgca gcacagagag 3720
ccagctagaa gcgctggccg cagagcagca gcccgggaac caggcccagg cccaggccca 3780
gctggccagc ctctactctg ccctgcagca ggccctgggg tctgtttgtg agagcaggcc 3840
tgagctgagt ggtgggggag actctgctcc ttccgtctgg ggccttgagc cagaccagaa 3900
tggagctagg agcctcttta agagagggcc cctgctgact gctctctccg ctgaggcagt 3960
agcatctgcc ctcctcaagc ttcatcaaga cctgtggaag actcaacaga cccgggatgt 4020
```

tctgagggat caggtccaga aactggaaga gcgtctaact gatactgagg ctgagaagag 4080
ccaggtccac acagagttgc aggatctgca gagacagctc tcccagaatc aggaagagaa 4140
atccaagtgg gaaggaaagc agaactccct agaatctgag ctgatggaac tacatgaaac 4200
tatggcatcc ttacagagtc gcctgcggag agcagagcta cagcgaatgg aagcccaggg 4260
tgagcgagag ttacttcagg cagccaagga gaacctgaca gcccaggtgg aacacctgca 4320
agcagctgtc gtagaagcca gggctcaggc aagtgctgct ggcatcctgg aagaagacct 4380
gagaacggct cgctcagcac tgaagctgaa aaatgaggaa gtagagagtg agcgtgagag 4440
agcccaggct ctgcaagagc agggcgaact gaaggtggcc caagggaagg ctgaagtgca 4500
gaatttggcc ctcctgaccc agaccctagc tgaaagagaa gaggaggtgg agactctgcg 4560
gggacaaatc caggaactgg agaagcaacg ggaaatgcag aaggctgctt tggaattgct 4620
gtctctggac ctgaagaaga ggaaccaaga ggtagatctg cagcaagaac agattcagga 4680
gctagagaag tgtaggtctg ttttagagca tctgcccatg gccgtccagg agcgagagca 4740
gaagctgact gtgcagaggg agcagatcag agagctcgag aaggatcggg agactcagag 4800
gaacgtcttg gagcatcagc ttctagaact tgagaagaaa gaccaaatga ttgagtccca 4860
gagaggacag gttcaggacc tgaaaaagca gttggttact ctggaatgcc tggccctgga 4920
actggaggaa aaccatcaca agatggagtg ccagcaaaaa ctgatcaagg agctggaggg 4980
ccagagggaa acccagagag tggctttgac ccaccttacg ctggacctag aagaaaggag 5040
ccaggagctg caggcacaaa gcagccagat ccatgacctg gagagccaca gcaccgttct 5100
ggcaagagag ctgcaggaga gggaccagga ggtgaagtct cagcgagaac agatcgagga 5160
gctgcagagg cagaaagagc atctgactca ggatctcgag aggagagacc aggagctgat 5220
gctgcagaag gagaggattc aggttctcga ggatcagagg acccggcaga ccaagatcct 5280
ggaggaggac ctggaacaga tcaagctgtc cttgagagag cgaggccggg agctgaccac 5340
tcagaggcag ctgatgcagg aacgggcaga ggaagggaag ggcccaagta aagcacagcg 5400
cgggagccta gagcacatga agctgatcct gcgtgataag gagaaggagg tggaatgtca 5460
gcaggagcat atccatgaac tccaggagct caaagaccag ctggagcagc agctccaggg 5520
cctgcacagg aaggtaggtg agaccagcct cctcctgtcc cagcgagagc aggaaatagt 5580
ggtcctgcag cagcaactgc aggaagccag ggaacaaggg gagctgaagg agcagtcact 5640
tcagagtcaa ctggatgagg cccagagagc cctagcccag agggaccagg aactggaggc 5700
tctgcagcaa gaacagcagc aggcccaggg acaggaggag agggtgaagg aaaggcaga 5760
cgccctccag ggagctctgg agcaagccca tatgacactg aaggagcgtc atggagagct 5820
tcaggaccac aaggaacagg cacgaaggct ggaggaagag ctggcagtgg agggacggcg 5880
ggtccaggcc ctggaggagg tgctgggaga cctaagggct gagtctcggg aacaggagaa 5940
agctctgttg gccctccagc agcagtgtgc tgagcaggca caggagcatg aggtggagac 6000
cagggccctg caggacagct ggctgcaggc ccaggcagtg ctcaaggaac gggaccagga 6060
gctggaagct ctgcgggcag aaagtcagtc ctcccggcat caggaggagg ctgcccgggc 6120
ccgggctgag gctctgcagg aggcccttgg caaggctcat gctgccctgc aggggaaaga 6180
gcagcatctc ctcgagcagg cagaattgag ccgcagtctg gaggccagca ctgcaaccct 6240
gcaagcctcc ctggatgcct gccaggcaca cagtcggcag ctggaggagg ctctgaggat 6300
acaagaaggt gagatccagg accaggatct ccgataccag gaggatgtgc agcagctgca 6360
gcaggcactt gcccagaggg atgaagagct gagacatcag caggaacggg agcagctgct 6420
ggagaagtct ctggccaga gggtccaaga gaatatgatc caagagaagc agaatctggg 6480
gcaagagaga gaagaggagg agataagggg ccttcatcag agtgtaaggg agctacagct 6540
gactctagcc caaaaggaac aggagattct ggagctgagg gagacccagc aaaggaacaa 6600
cctggaagcc ttacccaca gccacaaaac ctccccaatg gaggaacaat ctctaaaact 6660
tgattcttta gagcccaggc tgcagcggga gctggagcgg ctacaggcag ccctgagaca 6720
gacagaagcc agggagattg agtggaggga gaaggcccag gacttggcac tctccctagc 6780
gcagaccaag gccagtgtca gcagtctgca ggaggtagcc atgttcctac aagcctctgt 6840
cctggagcgg gactcagaac agcaaaggct gcaggatgaa ctggagctca ccagacgggc 6900
tctggagaag gagcggctac acagcccagg tgcaaccagc acagcagaac tggggtccag 6960
aggggagcag ggtgtgcagc tgggagaggt gcttgaacac ctgcagcaag cagtggccag 7080
tggaatggag aagcagtcat ggagacaaag gcttgaacac ctgcagcaag cagtggccag 7080
gctggagatt gacaggagca ggctgcagcg ccacaatgtc cagctgcgga gtaccttgga 7140
gcaggtggag cgagaacgga ggaagctgaa gagggaggcc atgcgtgcgg cccaggcagg 7200
gtcctagag atcagcaagg ccacggcttc ttcacccaca cagcaggatg ggagaggaca 7260
gaagaactca gatgccaagt gtgtggctga actgcagaaa gaggtggtcc tgctgcaagc 7320
tcagctgact ttggagcgga agcagaagca ggactacatc acccgctcag cacagaccag 7380
ccgtgagcta gcaggcctgc accacagcct ctcacactca cttcttgccg tggccaggc 7440
ccctgaggcc actgtcctgg aggcagagac ccgcaggctg gatgagtccc tgactcaaag 7500
tctgacatcc ccagggccag tcctgctaca cccagccccc agcactaccc aagccgcctc 7560
caggtagcag ccacagccag gagcacacag acagaagact gtgtcatggg tcatggcccc 7620
tccgcacacc tacaggtttg ccaaaggaaa agccatggctc tgttaggcac ccaggagccc 7680
caggtcggcg ggtgttccca ggaagaggaa gtaaatctgc aaccctgggg aggaccccaa 7740
ctcacctggg aatgaggcaa attgcatttg cttgctccct atggaatcac ccagaggggt 7800
gccttgccct ggctgaggga catgtactgc ctctcatcta gaatttattt tcctagcact 7860
tcaccacctc cttcatcttc tccttcaaca ataaaccctg accaaatgat tcaaaaaaaa 7920

tataaaaaag accagtccaa gcttattccc tttagtgagg tt 7962

<210> 16
<211> 1373
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte Clone No: 2195418

<400> 16
cagccgggca gccgcttccc gcccccgagc aggagccggt gcgagcggag cagagccgag 60
gtcgggccgc gagcggagcc ggctgagcgg gcgccgagct cccgccatgg cccggaacac 120
gctgtcctcg cgcttccgcc gggtggacat cgacgaattt gacgagaaca aatttgtgga 180
cgagcaggag gaggcggcgg cggcggcggc ggagccaggc ccggacccga gcgaggtgga 240
cgggctcctg cggcaagggg acatgcttcg ggcattccat gcagccttgc ggaactctcc 300
cgtcaacacc aagaatcaag ctgtgaagga gcgagcccag ggcgtggtgc tgaaagtgct 360
cacaaacttc aagagcagtg agattgagca ggctgtgcag tcactggaca gaaacggcgt 420
tgacttgtta atgaagtaca tttataaagg ctttgagaag cccacagaaa atagcagcgc 480
agtgttactc cagtggcacg aaaaggcctt agcagtagga ggactaggct ccattataag 540
agttcttaca gcaagaaaga ctgtttaaaa aaaataaaaa gactcatgtt accttgagaa 600
gaattctgga tgcccaggct ggtgaagaag ggattgacaa tggaccatct tcctaggaac 660
tcccaagtaa actatttcag gacatgtatc tgctgaaatg tattttattt tcaaggtgga 720
ggggaaaatc gtctgtttcc taaatcctgt ttaggatctg atagtctatg cctttgtctc 780
cgagtactgc agaactgaca ttttgacggt ctaccagcgt ggcggctggt gttggtcaga 840
tgcacctgtg tgcactgggg gagggatggt ttgggcaggt gcagatccaa gggctgtggt 900
aaacgggaga gcttgtgttt ttgaagtgga aaaaaaccca agagtttgta cagacatcct 960
gtcttcccag agaaggtgga cactcttggg ctcattgtaa agtgcctgct gcatcaataa 1020
agctcttggc ttattagtct atacattgcg gtgtgtttcg tgtatgtaaa aaaaaatggt 1080
aatgaatggg atggtaatga atgagagttc agttgttgtt ccggaaaccc gatgtggaag 1140
gagtagacct gtgtccctgt tgagccaccc ctgggagcga gcatggcaat ccacaggccc 1200
tctgccacag gacgccagcc tcggcctcag agctgccggc tgctgcagag aggtgtttgc 1260
tgaataaact atttattgtt tcttattcct ttgatttgta tgtaattaat tttggagctt 1320
atttaattaa tttaataaag tgccaaacat ttaataatta aaaaaaaaaa aaa 1373

<210> 17
<211> 151
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> GenBank ID No: g2282042

<400> 17
Met Ser Lys Asn Thr Val Ser Ser Ala Arg Phe Arg Lys Val Asp
1 5 10 15
Val Asp Glu Tyr Asp Glu Asn Lys Phe Val Asp Glu Glu Asp Gly
20 25 30
Gly Asp Gly Gln Ala Gly Pro Asp Glu Gly Glu Val Asp Ser Cys
35 40 45
Leu Arg Gln Gly Asn Met Thr Ala Ala Leu Gln Ala Ala Leu Lys
50 55 60
Asn Pro Pro Ile Asn Thr Lys Ser Gln Ala Val Lys Asp Arg Ala
65 70 75
Gly Ser Ile Val Leu Lys Val Leu Ile Ser Phe Lys Ala Asn Asp
80 85 90
Ile Glu Lys Ala Val Gln Ser Leu Asp Lys Asn Gly Val Asp Leu
95 100 105
Leu Met Lys Tyr Ile Tyr Lys Gly Phe Glu Ser Pro Ser Asp Asn

```
                110                 115                 120
Ser Ser Ala Met Leu Leu Gln Trp His Glu Lys Ala Leu Ala Ala
                125                 130                 135
Gly Gly Val Gly Ser Ile Val Arg Val Leu Thr Ala Arg Lys Thr
                140                 145                 150
Val
```